U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE
NATIONAL METEOROLOGICAL CENTER

OFFICE NOTE 215

How to Pick Another New Forecast Model (Spectral vs 7L PE)

John D. Stackpole
Development Division

JULY 1980

Contents

I. Introduction

During the latter portions of 1979 and the first half of 1980 the Systems Evaluation Branch of NMC's Development Division (NWS/NOAA) undertook the "preimplementation" evaluation of a 12 layer Primitive Equation model with a spectral representation of forecast variables. The intent of the test was to determine whether the Spectral model was a worthy replacement for the then operational 7L PE model. It was, although the testing brought out a number of deficiencies which had to be corrected (and re-tested) prior to the operational implementation.

What follows is a brief description of the models as initially compared, the test procedure, the results (in part), the alterations to the Spectral model and the results of the final retest.

II. The Contending Models

A. 7L PE

This is basically the NMC operational large scale forecast model. Originally it was a 6 layer model on a 381 km mesh (Shuman & Hovermale, 1968); a series of tests (Stackpole et al, 1978) lead to the implementation of a 7 layer PE model on a 190.5 km grid (a half-bedient mesh); these and subsequent developmental changes to the 7L PE model are detailed in NMC Office Note 177 (Stackpole, 1978).

Since the introduction of the above changes additional modifications have been made in four broad categories - i) the "long wave cooling" has been adjusted to cause the available potential energy of the model to remain nearly constant in time; ii) a sigma-coordinate energy conserving formulation of the hydrostatic equation has been introduced; iii) the horizontal diffusion has been changed (to a non-linear form) so as to assure approximate conservation of the model's total kinetic energy; iv) the convective precipitation parameterization has been somewhat generalized. Changes i), ii), & iii) are detailed in Cooley, 1979 (b); change iv) is detailed in Cooley, 1979 (a) in the context of the LFM-II model - the same changes were introduced into the 7L PE in October 1979.

In an effort to stone two birds, further changes were made to the 7L PE model for the comparison tests. These changes, aimed particularly at improving the tropical regions of the 7L PE forecasts, were judged, *a priori*, to have little likelihood of altering the forecasts in the more northernly extent of the grid. Thus the comparisons with the Spectral model in those regions would still be valid with respect to the operational version of the 7L PE model. For the most part this judgement proved correct with two exceptions. These are detailed further on.

The tropical region changes were twofold: i) the initialization section of the model was modified to accept actual analyses south of 9° N. Previously a meteorologically bland set of fields were placed there (and fared into the real analysis further north) to promote stability.

Appropriate adjustments to the coriolis parameter, map factor, and surface elevation fields were made in conjunction with the introduction of the real analyses. ii) In the forecast section, the horizontal boundary conditions were changed from a free slip wall (located between the penultimate and ultimate grid point rows or columns) to a "diffusive nudge" (or "mushy") boundary condition. In this the outer row quantities retain their initial values throughout the forecast; the values at rows 2, 3, 4, and 5 are nudged toward their initial values, ϕ_i, by adding

$$K\nabla^2\left(\phi^{\tau-1} - \phi_i\right)$$

to the past time step values, $\phi^{\tau-1}$, each time step of the forecast. K has a fixed value of 0.04 (dimensionless) and ∇^2 is a finite-difference 5 point (+) laplacian operator. ϕ represents all of the forecast variables: u, v, θ, p_σ , and q.

It was in the context of the tropical modifications that two related errors arose, unfortunately, and remained undetected until the test series was completed. The first was that the outer row initial condition for the specific humidity was set (and of course remained so throughout the forecast) at the equivalent of 2% relative humidity. The interior rows had values appropriate to the analysis in the tropics - considerably higher values. The second error, which compounded the first, was that the specific humidity was not included in the diffusive nudge or mushy boundary condition calculation at all. Thus the numerical effect of a very sharp humidity gradient could, and did, propagate inward from the edges of the grid, resulting in a rather raged looking humidity forecast. This was remarked upon at the time but erroniously attributed to another cause.

This other cause was the second discrepancy between the operational and test versions of the 7L PE that could have an impact in extratropical regions. Again inadvertently, the test version of the code had a somewhat lighter non-linear diffusion than the operational version. The effect of this appears to be small and limited to the wind speed forecasts; indeed it seems to have been slightly beneficial - the wind speed forecasts (particularly in jet maxima) were slightly better in the test 7L PE than the operational. But this had to be kept in mind in the comparison of the 7L PE and Spectral models.

B. 12 L Spectral

This model, the contender to become the operational large scale forecast model, is fully described by Sela, (1980). In outline it is a 12 layer primitive equation hydrostatic forecast model employing a global or hemispheric spherical harmonic representation of the variables in the quasi-horizontal surfaces and layers of the sigma coordinate system. For the initial round of comparison tests the twelve layers were equally spaced in pressure from the earth's surface to zero pressure. For the first portion of each forecast (more on this anon) the spectral truncation was rhomboidal with 30 modes and a global extent; for the latter portions the truncation was hemispheric with rhomboidal truncation at wave 24. The time step (again for the first set of comparisons) was "semi-implicit backward", a system with a fair degree of damping inherent in it, and there was no "initialization" of the initial conditions beyond that which is implicit in the Hough analysis used to supply those initial conditions. The "physics package" of the model is relatively complete: it includes orography, surface drag, heating and evaporation from the sea surface (but not from land), large and convective scale precipitation parameterizations, and horizontal (but not vertical) frictional dissipation. At present there is no inclusion of any of the thermal effects of radiation or, obviously, the diurnal cycle.

Sela (1980) describes the extensive development and testing which lead to the judgement that the model was indeed a worthy contender for the large scale operational forecast chore at NMC.

C. The Test Configuration

One of the constraints on the comparison tests was that the models would be run in an operational configuration. For the 7L PE this meant rerunning the test cases in the current operational configuration - fine mesh (1/2 bedient or 190.5 km mesh), second order finite differences to 60 hours, then (for the 00Z initial time cases) coarse mesh (1 bedient mesh) and 4th order finite differences to 144 hours. See Campana (1978) for details on the 4th order system. The switchover from fine to coarse mesh is painless - the appropriate codes merely select every other gridpoint from the fine mesh and proceed. A small amount of noise is generated in the process - it quickly damps out, and no harm is done.

The potentially operational configuration for the Spectral model was selected on the basis of running times and meteorological desirability. The selected configuration was global with 30 modes (G3012) to 48 hours followed by hemispheric 24 mode (H2412) on out to 144 hours. The switchover from G3012 to H2412 is a touch more cumbersome than for the 7L PE - it is necessary to strip away both the high resolution modes and the low resolution ones that are uniquely global in character. This too introduces noise which seems to be well controlled by the horizontal diffusion and the damping character of the backward time step.

Both models start from the same analysis - the operational Hough analysis in coefficient form. The 7L PE proceeds in the usual way, generating mandatory level grid point analyses from the Hough coefficients then interpolating to σ coordinates, etc. The spectral system converts the Hough coefficients analytically to the appropriate spherical harmonics and then proceeds to the σ coordinate interpolation.

Table 1 presents a summary of the salient differences between the two models, as they were run in the first round of testing.

Characteristic	7L PE	SPECTRAL
Layers	Boundary; 3 tropospheric; 3 stratospheric;	12 equal in p
Horizontal Representation & Forecast length	1/2 bedient, 2nd order to 60 hr.; 1 bedient, 4th order to 144 hr.	G3012 to 48 hr.; H2412 to 144 hr.
Initial Analysis	Hough	Hough
Initialization	No	No
Time Step	10 or 20 min Leapfrog with "pressure gradient averaging"	17 or 34 min implicit backward
Radiation	Simple	None
Precipitation	Moisture in 3 lowest layers	Moisture in 8 lowest layers
Boundary Energy Exchange	Ocean Warming; land warming	Ocean Warming
Surface friction & Orography	Yes	Yes
Horizontal Diffusion	Non Linear $[\nabla^2() * \|\nabla^2()\|]$ with coefficient less than operational	Linear $[\nabla^4()]$
Tropics	Revised initial conditions; Diffusive nudge boundary conditions	Nothing extra beyond global domain

Table 1 Model Salient Features

III. Experimental Design

A. Uses of the Operational Model

In that the intent of the comparison testing was to certify the appropriateness of the Spectral model becoming the Operational Forecast Model at NMC, particular, but not exclusive, attention was paid to those portions of the model of particular operational utility. In broad categories they are:

1) Aviation Forecasting

Here the concern is for the quality of the short range (24-30 hour) forecasts of pressure and wind patterns as they relate to general and commercial aviation flight planning. Of particular interest are the strength and location of the jet stream, the tropopause pressure and wind shear, the temperature fields at 250 and 100 mb., and to a lesser extent, the synoptic patterns at lower levels.

2) Marine Forecasting

Here the primary concern is for the sea level pressure forecasts, and the 500 mb as it impacts the interpretation of the SLP maps, at 24 and 48 hours, in the Western Atlantic, Eastern Pacific and Alaskan regions.

3) Quantitative Precipitation

Although the LFM-II model supplies the routine guidance for the U. S. area in this and other categories, there is still a continuing interest in seeing if the rain forecasts can be improved by any new model. The concern is for both areal coverage of rain/no rain indications and quantitative amounts.

4) General Quality and Crediability

This is rather a catch-all category - a new forecast model should produce forecasts that "look right", have vertical consistency, reasonable height, isotach, and vorticity patterns and not be wildly in error in special cases.

5) Tropical Forecasts

NMC has a requirement for tropical and southern hemisphere forecasts for aviation which has been met in the past by a blend of analyses (as persistence forecasts) and dynamic forecasts. Now the Spectral model introduces the possibility of using forecasts everywhere. The spectral tropical forecasts were given special evaluation, in terms of the surface and 250 mb stream function forecasts. There was no comparison with the 7L PE model in this case.

6) Medium Range Forecasts

The operational guidance for the five day forecast program includes many items - for the purposes of the intercomparison three were concentrated upon: the 84 hr. ("day 3") sea level forecast, the 5 day mean and departure from normal 500 mb heights centered on day 3, and 24 hr. and 5 day accumulated precipitation, also centered on day 3. The medium range forecast comparisons took place for only those cases having a 00Z initial analysis time.

7) Model Output Statistics (MOS) Forecasts

Most of the MOS forecasts take their input from the LFM-II model; however the forecasts for Alaska and some medium range forecasts for the Pacific North West come from the Operational model. The appropriate fields needed for the various MOS forecasts were included in the output from the Spectral model and extensive intercomparisons were undertaken by the Techniques Development Laboratory (the progenitor of MOS) of the NWS.

B. Selection of Cases

In an ideal world the selection of test cases would involve simply turning to a library of historical situations and picking out those that exemplified the particular uses of the forecasts outlined above. Unfortunately, this was not possible. The library situations were generally saved on the basis of continental U. S. meteorological phenomena: big storms and other special situations or else (sometimes "and") because one of the operational models showed a particularly bad error. The cases then were selected without much attention to their extra-U. S. characteristics but as many as possible of the various usage criteria that could be applied were applied in the selection process. The unstated assumption was that what was good for the U. S. would be good for the rest of the Northern Hemisphere.

The selected cases, and an indication of why they caught the eye of the selection committee, were:

1. 12Z 4 February 1978:

   The great New England Blizzard of '78. This is the one that the LFM did real well on (at 48 hours) and the 7L PE (run after the fact - there was a power outage that weekend) failed to predict with the same degree of assurance.

2. 00Z 16 January 1979:

   Atlantic Blocking.

3. 12Z 12 February 1979:
   Overdevelopment forecast (by the 7L PE) in the lee of the Rockies - a longstanding 7L PE problem.

4. 12Z 2 March 1979:

   Jet stream cross contour flow in the Asian - North Pacific Region.

5. 00Z 11 April 1979:

   This is a sample from a series of poor sea level and precipitation forecasts selected by the Mid-Latitude Storm Committee member on the selection group.

6. 12Z 9 May 1979:

   A locked in error case (the only one all last winter and spring).

7. 00Z 30 August 1979:

   A case selected by the 5 day forecast people of particular interest to them - it involved a substantial change in mean flow patterns.

8. 00Z 5 September 1979:

   Hurricane David over the South Eastern U. S.

9. 00Z 30 September 1979:

   A poor precipitation forecast over the upper & mid-Mississippi Valley and western upper Great Lakes region with a developing wave.

10. 12Z 8 October 1979:

   The great Washington, D. C. blizzard of '79.

Obviously not all of the cases selected dovetail with the operational utility criteria one to one - the (implicit) hope is that there is sufficient variety in these ten cases to cover the major areas of interest.

C. Forecast Evaluation

i. Subjective

Even though the cases selected may not have matched the operational utility criteria very closely, the subjective evaluations did. Particular NMC Forecast Division forecasters with primary forecase responsibilities in the various utility categories were called upon to consider the two forecasts from each case and make qualitative judgements of them. Questionnaires (Appendix I) were prepared to aid the forecasters and to focus their attention. The specialists and the sets of maps they received were:

1. Aviation. Roy McCarter

   24 hr sea level pressure & 1000-500 mb thickness
   24 hr 500 mb heights & vorticities
   24 hr 250 mb heights, isotachs & temperatures
   24 hr tropopause pressure & vertical wind shear
   24 hr 100 mb heights, isotachs & temperatures

2. Marine forecasting. Harry Brown

   24 & 48 hr sea level pressure & thickness
   24 & 48 hr 500 mb heights & vorticities

3. Precipitation forecasting. Dave Olson

   24, 48 & 84 hr sea level pressure & thickness
   24, 48 & 84 hr 500 mb heights & vorticity
   24, 48 & 84 hr mean relative humidity & 700 mb vertical velocities
   24, 48 & 84 hr precipitation (Varian & hand trace)

4. General Quality. Harlan Saylor

   All of the above

5. Tropical. Ed Carlstead

   24 & 48 hr 1000 mb streams
   24 & 48 hr 250 mb streams

6. Medium Range Forecasts. Fran Hughes & James O'Connor

   84 hr sea level pressure & thickness
   84 hr 500 mb heights & vorticity
   D+3 5-day mean 500 mb heights & departure from normal
   D+3 24 hr accumulated precipitation centered at 84 hours (100 station list)
   D+3 5-day accumulated precipitation centered at 84 hours (100 station list)

These maps were, for the most part, the usual Varian contour maps that the forecasters use routinely: $1{:}30\times10^6$ polar stereographic projections of the Northern hemisphere. One of the exceptions was the stream function maps - they were "Mercator strip" projections extending from 60°N to 60° South. (The requirement for information to 60°S precluded the preparation of these maps from the 7L PE model).

The other non-standard map was the precipitation. Since Varian mapping codes work with the 381 km $1{:}30\times10^6$ scale polar stereographic map projection grid, the final output from each of the contending models had to be interpolated from the individual model's forecasting grid to that grid. The precipitation forecasts, being discontinuous, would suffer considerable debasement in the interpolation process and a different procedure was followed: The models' precipitation forecast were printed on the scale of the particular model's grid, transcribed to a suitable map base and hand analyzed. This map thus contained all of the resolution that the models were capable of, and was the map that went to the forecasters. The Varian precipitation maps were also produced for reference as they would be the operationally generated maps.

Along with the rather sizeable number of forecast maps, each forecaster received an appropriate set of observed analyses for verification purposes. Each case involved the preparation of some 75 original maps plus the copies for each forecaster. There is always the fear in such evaluations that we will overlook something important.

ii. Objective

Of necessity the objective verifications in the form of various error measuring statistics cannot be tailored uniquely to the model usage criteria; instead we used some standard verification methods and programs, some portions of which can be interpreted as relating to portions of the usage criteria.

For the objective verifications a standard set of statistics was calculated, to wit: mean (bias) and root mean square (rms) errors of geopotential heights, temperatures, relative humidity and wind speeds, plus rms vector wind error, Tewles/Wobus S1 Score (Brown, 1971), and threat score and bias of precipitation forecasts. The forecasts were verified for 24, 48 and 84 hours (although not all the cases were verified for all the time periods because of missing verification data) and at the 100 kPa, 85 kPa, 50 kPa, 25 kPa and 10 kPa mandatory pressure levels.

The data against which the forecasts were tested were of two kinds: gridded analyses (The NMC FINAL analyses) and radiosonde upper air observations. (Raob measurements at 100 kPa were not used, thus avoiding problems introduced by various "reduction-to-sea-level" methods when the 100 kPa surface was underground). Based on previous studies (Stackpole, *et al*, 1978) we can assert that the verifications-against-analyses and the verifications-against-observations lead to the same conclusions; in the interest of reducing bulk somewhat, the statistics incorporated in this report are those of the verifications against observations only (except at 100 kPa where the analyses were used).

Three networks of observation stations were used: 110 stations over North America (essentially all the regularly reporting Raob stations from 25° to 60° North latitude and 50° to 145° West longitude); 102 stations quasi-uniformly distributed over the entire Northern Hemisphere; and 93 stations in the tropics between the equator and 30°N.

The method of calculation of the various error statistics is straightforward with one exception, the S1 score. For the mean and rms error statistics the forecast quantities were biquadratically interpolated to the station locations, the errors established and the appropriate summations over all the stations with valid reports in the network performed.

For the S1 score calculation, a preliminary pass is made through all of the available upper air observations (not just those of the network in question) and the station which is the nearest neighbor to each of the network stations is located. Then the observed and forecast height gradients between the station pairs are used for the S1 score calculation. The "nearest neighbor" selection is limited by claustrophobic (pairs closer than 100 km are not allowed), agoraphobic (pairs separated by more than 2357 km are not allowed) and geminiphobic (if A selects B as its closest neighbor, B may not select A) constraints. This method of calculation of S1 differs from the usual one in which the gradients are computed between pre-selected grid points in a fixed geographic array. Again comparison between the station S1 and grid S1 scores for the various forecasts and models showed no signficant differences in the conclusions one would draw from them. Excepting 100 kPa, station S1 scores are presented below.

For the objective verification of precipitation forecasts a different network of 60 first order stations (long in use by NMC Forecast Division) was augmented by 30 additional stations designed to fill some gaps and cover problem areas in coastal and mountain areas. A computer algorithm was readied, designed to interpolate (in a manner appropriate to the discontinuous precipitation fields) from the grid points at which precipitation was forecast in the models to these stations. Each of the models incorporated this computation in their output sections, thus producing a list of 12 hour accumulated precipitation amounts for the verification times and stations. These station forecasts were the material for the calculation of the precipitation threat and bias scores.

The medium range forecasts were objectively verified (by the forecasters concerned) with statistics appropriate to the different nature of the forecasts - pattern correlations for the departure from normal maps and Heidke Skill scores (relative to climatology) for the accumulated precipitation.

IV. Results - First Round

A. Execution of the Tests

No particular difficulty was experienced in running the forecast models on the test cases, other than the usual kinds of problems generally associated with large scale computers. A fair amount of mickey-mouse chores were necessary - establishing data sets, designing job flow sequences, archiving intermediate results, making originals and copies of maps and, as ever, rerunning jobs when "something went wrong". A number of delays were encountered in getting the 7L PE, with the modified boundary conditions, underway - once overcome however, things went smoothly (with the exception of the difficulties alluded to above).

B. Subjective Evaluations

The results of the subjective evaluations are most conveniently presented as tabulations of the responses to the questionnaire, with an arbitrary scoring system, and a reporting of comments made by the evaluators both on the questionnaire forms and during subsequent discussions. The scoring system is simply to assign one point to whichever model was preferred in each subcategory on the questionnaire plus an additional half-point if the preferred model was "much" better than the other. The results are arranged here by field of specialization.

i. Aviation Forecasting (McCarter)

Category	Preference Tally		
		24 Hr	
250 mb Winds	SPEC	7L PE	Ties
Atlantic	3	4	3
N. America	1	7	2
Pacific	2	4	4
Sub-Total	6	15	9
250 mb hts. & temps			
Atlantic	0	10	0
N. America	1	5.5	4
Pacific	0	6.5	4
Sub-Total	1	22	8
Tropopause & Shear	4	2	4
100 mb hts.	1	3	6
100 mb winds	0	4	6
100 mb temps	4	5	1
General Aviation	0	2	8
Grand Total	16	53.5	42

Table 2

Interpretation of the tally in Table 2 is not too difficult: a consistent preference for the 7L PE except for the tropopause pressure and tropopause vertical wind shear. In only three instances however was the 7L PE given a "much better" half point bonus - the model differences were small, although consistent.

The evaluator's comments pinpointed the problems. The jet stream (as shown in the 250 mb isotachs) was quite consistently weaker in the Spectral model relative to the 7L PE. (They were both weak relative to the verifications but the Spectral was more so). The 250 mb temperature in the Spectral showed a warm bias (over that of the 7L PE) and this was the usual reason for the selection of the 7L PE as better. Some of the ties came about because both forecasts were poor; there was little point in deciding which forecast was less dreadful.

In general, however, the evaluator concluded that the differences between forecasts were small and that he (somewhat surprisingly) would not object to implementation of the Spectral model as it was.

ii. Marine Forecasting (Brown)

Category	24 HR			48 HR			24 & 48 HR Combined	
	SPEC	7L PE	Ties	SPEC	7L PE	Ties	SPEC	7L PE
SLP & Thickness								
Atlantic	2	5	3	2	7.5	1	4	12.5
Pacific	6	0	4	2	6	2	8	6
Subtotal	8	5	7	4	13.5	3	12	18.5
500 mb Ht & Vorticity								
Atlantic	3	2	5	1	4	5	4	6
Pacific	5	3	2	2	4	4	7	7
Subtotal	8	5	7	3	8	9	11	13
Grand Total	16	10	14	7	21.5	12	23	31.5

Table 3

The results from this tally are somewhat more mixed than those of the previous one: the short range forecasts give the spectral a slight edge, mainly based on the Pacific area; the scales tip the other way at the two day range, sufficiently that the combined grand total of points favors the 7L PE. The SLP comparisons showed fewer ties and were the larger contributor to the 48 hr preference for the 7L PE. Only one forecast category got a "much better" bonus.

The evaluator was rather exiguous with his comments but one might be considered significant: after attempting a second look at the comparisons he found he was not able to reproduce his initial judgments with any confidence - differences that had seemed significant first time through no longer seemed so and others had taken their place. In effect, the models are essentially tied as far as SLP marine forecasting concerns go.

iii. Quantitative Precipitation Forecasting (Olson)

Category	24 HR			48 HR			84 HR			24, 48 & 84 HRS	
	SPEC	7L PE	Ties	SPEC	7L PE	Ties	SPEC	7L PE	Ties	SPEC	7L PE
Rain/No Rain Coverage											
East	1	8.5	1	3	7.5	1	1	2	2	5	18
West	1	5.5	4	4	4	2	2	3.5		7	13
Subtotal	2	14	5	7	11.5	3	3	5.5	2	12	31
Quantitative											
East	3.5	6.5	1	1.5	5.5	4	0	3.5	2	5	15.5
West	1	6.5	3	3	6	2	2	2	1	6	14.5
Subtotal	4.5	13	4	4.5	11.5	6	2	5.5	3	11	30
Total	6.5	27	9	11.5	23	9	5	11	5	23	61

Table 4

I have not bothered to tabulate the Relative Humidity pattern preferences - they favor the spectral model almost 100%. This preference is due almost completely to the excessive roughness of the R. H. patterns in the 7L PE model which, in turn, is due to the combination of the horizontal boundary condition error and reduced smoothing alluded to previously.

The "Utility of Forecasts" section is also untabulated - the two models received equal scores - it is little more than a confirmation of the judgments made in the Marine Forecasting section.

The tabulations shown in Table 4 allows for little variation in interpretation - the 7L PE has the clear preference in all categories and at all time ranges.

The comments include the familiar observation that the model forecasts frequently resemble each other more than the atmosphere and that in critical QPF cases the 7L PE did a "little better" but that neither model was really desirable. A more telling comment was the observation that since the LFM-II is the primary guidance model for the operational precipitation forecasts it doesn't really matter too much what the Spectral or 7L PE have to say. This is a slightly defeatist (although realistic) viewpoint, but still one would like to see improvements in a new model.

iv. General Quality & Credability (Saylor)

The evaluator in this category made use of the questionnaire forms but a tabulation of them shows no significant differences from the other expert's judgments. His comments also confirm those of his fellows: the models are "remarkably similar" making good, and bad, forecasts together; the jets are too weak by 5-10 mps; the tropopause vertical wind-shear was much improved in the spectral; the spectral precipitation was poorer than the 7L PE but that's not so important; the spectral was particularly lacking in convective precipitation. This evaluator saw no bar to immediate implementation.

v. Tropical & Southern Hemisphere (Carlstead)

Here there was no preferential tabulation as only the spectral forecasts were available on the Mercator (60°N to 60°S) projection maps. The evaluator's observations and comments were directed toward the question of the prospective value of the model in the tropics and southern hemisphere, in an absolute sence. His comments were very extensive and thorough - his general conclusion was that the model "handled the Tropics well, better than I had expected". A couple of problem areas were noted: the south-east trades in the area from 0-20°S latitude and 100 -160°W longitude were consistently weak and mal-formed; the Argentine area forecasts were quite unsatisfactory (the evaluator speculated, probably correctly, that this latter problem arose from an inadequate spectral representation of the Andes Mountains). A number of common themes recurred in his case-by-case evaluations: the jet stream winds were too weak; monsoons seem to be well forecast; Typhoons and hurricanes are captured, at least in a large scale sence. The evaluator devised his own scoring system (1 - useless, 2 - poor, 3 - adequate, 4 - good, 5 - excellent) and the all-case average score was 4.2 ± 0.42, well in the "good" range.

The medium range and MOS forecasts were evaluated entirely objectively and are discussed in the next section.

C. Objective Verifications

i. Heights, Winds & Temperatures

The objective verifications of the dynamic forecast quantities (heights, temperatures, winds) are of necessity of a more broad brush nature than the specific use orientation of the subjective evaluations. They also present something of a problem in presentation: since many (5) levels, many (6) areas, many (5) variables, many (3) forecast verification times and many (3) statistics were computed for each of the 10 cases, some selection has to be made. Averaging over the cases is not appropriate as many of the verification statistics have a normal seasonal variation that could mask the inter-model variations. Large (even small) tables of numbers are at best a considerable barrier to ease of understanding - the best compromise is a selection of scatter diagrams of Spectral model verification statistics plotted against 7L PE statistics for all the cases. These afford a reasonably easy way of assessing whether there are appreciable differences between the models and, if we look at enough of them, do give a fairly reliable indication of the overall relative quality of the two models.

The selection of which statistics, levels, etc. to present was based on a combination of tradition (500 mb and sea level S1 scores are mandatory in any NMC verification), consciousness of the primary uses of the model forecasts, and limitations of bulk.

In the diagrams the ordinate is the Spectral model verification statistic, the abscissa the 7L PE; this means that (for positive definite error statistics - rms errors, S1 score) points falling to the right of the 45° line indicate the Spectral to be better than the 7L PE. For signed errors (bias) the rule of which is better is not quite so simple - but it will be obvious in each case.

The first group of diagrams (Figures 1 - 8) are for the NH102 network of 102 radiosonde stations spread quasi-uniformly over the Northern Hemisphere north of 30° latitude. The first of these, Figure 1, shows the 500 mb S1 for all the available cases and forecast hours (as all of the figures will), and also clearly shows the models to be of equal quality (as not all of the figures will). The words "SPECTRAL BETTER" are not to be taken as a judgment of relative merit but merely as a reminder of the region of the chart in which "Spectral model better" points will fall. Figure 2 shows the rms height error in meters (S.I. units are used throughout) and suggests a marginal improvement of the Spectral over the 7L PE. The improvement is rather small (10 meters) and probably not meteorologically significant. Figure 3 is a sort of a callibration of Figure 2, and a verification of my earlier assertion that verifications-against-analyses will lead to the same conclusions as verifications-against-observations. It shows the same statistic as Figure 2 but the verifying data are the gridpoint values of the NMC analysis over the area of the old NMC Octagon. The agreement as to conclusions between the figures is apparent. (The slightly lower value for the error numbers in Figure 3 doubtless arises from the smoothness of the gridpoint analysis relative to the directly observed heights).

Figure 4 takes us up to 250 mb for the rms height error. Clearly a tie. (The 250 mb S1, not shown, also indicated a tie). Figure 5, the 250 mb mean temperature error, or bias, is the first initimation of problems: every point falls in the "7L PE better" region between 45° and 135° - both models are too warm, on the average, with the Spectral averaging about one degree (celcius) warmer than the 7L PE. This is not a large difference, but the consistency is a matter for concern, and the errors have a bearing on aviation forecasting. The subjective evaluator noted this warm bias in his case-by-case studies. Figure 6 exhibits the same preference for the 7L PE model, this time in terms of the rms temperature error.

The 250 mb wind speed errors are exhibited in Figure 7. The results confirm the subjective evaluator's comments that the jet stream maxima are too slow and indeed make it clear that the winds are forecast too slow (in both models) in general. And further, the spectral is even slower than the 7L PE. The rms vector wind error, Figure 8, by indicating the models to be of equal quality, suggests that the Spectral is doing as good as, or better than, the 7L PE in forecasting the wind patterns, sufficiently so to overcome the speed bias error.

The next group of figures (9 - 20) are verifications in the North American area - all of the available radiosondes between 25° to 60° North and 50° to 145° West are used, or, in the case of Figure 9 and 11, the analysis. Figure 9 shows the sea level S1 score, computed from the analysis, using the customary 49 point grid. Another tie. Figures 10 and 11 are companions - both are the 500 mb S1 score, one with respect to the observations, the other against the analysis. Once again they both show the same: that the two models are equal in skill in S1 terms. The 500 mb rms height error, Figure 12, and the same statistic at 250 mb, Figure 13, do not alter the conclusion.

On the other hand, the mean and rms temperature errors, Figures 14 and 15, indicate that the North American area is not exempt from the hemispheric error characteristics noted in relation to figures 5 and 6: both models are too warm and the spectral model is more so. Similarly Figure 16 shows the two models both forecasting wind speeds too slow (for the most part) and, generally, shows the spectral model slower than the 7L PE.

Moving up some more to 100 mb (still using the North American station network) brings out some other, less conclusive, results. Figure 17, the temperature bias, shows rather a wide dispursion of points - on the average the models are about equal in quality (both about 1.5° too warm), but the dispursion suggests that it would be a better conclusion to say that both are equally poor. Figure 18, the rms temperature error, confirms this and, further, suggests that the Spectral is more poor than the 7L PE. Figures 19 and 20, the mean and rms wind errors, further the impression that neither model is particularly good at the 100 mb level (a fact long recognized for the 7L PE) and that the Spectral has not improved matters over the 7L PE. Quite the contrary. A particular item to note on Figure 19 is the positive bias in the Spectral wind speeds forecast as opposed to the slightly negative or nearly zero (average) bias for the 7L PE. This is in rather striking contrast to the 250 mb level results.

ii. Precipitation

Extensive case by case evaluation of the precipitation forecasts, verified in terms of observed and forecast amounts at networks of stations, has been undertaken by R. Hirano. Figures 21 and 22 serve to summarize the results most concisely. Figure 21 which speaks to the coverage (rain/no rain) aspects of the forecasts shows, on the top portions of the figure, the precipitation threat score (at 12 hour intervals), for both western and eastern networks of observations. These scores are for all the cases combined. The impression gained from these scores (and from consideration of the individual cases) is that the two models are about equal in their forecasting skill, or lack thereof. This is in contrast to the subjective evaluator who rather favored the 7L PE - what this mainly shows is that the threat score is, by itself, not a precise measure of what a subjective evaluator thinks is important.

The lower portion of Figure 21 indicates a more substantial difference between models: what is plotted is the count of the total number of stations, for all ten cases, for which rain was either reported (the column-bar graph) or forecast (x - 7L PE; . - Spectral) in each 12 hour period. These numbers are the ingredients of the usual bias score. They show clearly that the Spectral model forecasts less precipitation coverage than the 7L PE. Over the west this appears to be a good thing as the 7L PE (as has been its wont for many years) overforecasts there; in the east it appears to be too much of a good thing. Case-by-case study showed that one of the principal contributors to the low bias for the Spectral in the east was that model's almost complete failure to generate convective rain at appropriate places and times. This in spite of the model having a rather sophisticated convective parameterization.

Figure 22 summarizes the quantitative aspects of the forecasts. The upper pair of graphs are a quantitative threat score. This score is defined in the same manner as the customary rain/no rain threat:

$$T = \frac{H}{F+O-H}$$

except that F, O & H no longer stand for a simple count of the number of forecast, observed and correctly forecast (hits) precipitation points. Instead they represent the sum over stations of the actual quantitative amounts forecast and observed. H, the correctly forecast amount, is the sum of the lesser of F or O at each reporting station. The quantitative threat shows a consistent preference for the 7L PE.

The lower quartet of graphs round out the quantitative study. They show the total amounts of rain observed and forecast, summed over all stations, and the average amount per station. These last indicate a curious phenominon - while the Spectral model underforecasts the number of precipitating stations (Figure 21) in the East, it puts out too much rain at those (too few) stations. The 7L PE does the reverse in the first 48 hours - it forecasts too little rain at too many stations. Curious, and difficult to resolve.

iii. Medium Range Forecasts

Statistical evaluations of the two models' forecasts at the medium range (5 days) were undertaken by F. Hughes of NMC's Forecast Division. He couched his verifications in terms of the statistics routinely used for verifications of the medium range forecast guidance: pattern correlations of the "Day 3" (84 hour) sea level pressure forecast with the verifying analysis; correlations of the 5 Day mean departure from normal (DN) 500 mb height pattern centered on Day 3 with its verification; and Heidke Skill Scores (relative to climatogy) of the 24 hour and 5 day accumulated precipitation amounts both centered on Day 3. Table 5 exhibits the pattern correlations over North American and the U. S. subarea for the five forecasts run from a 00Z initial time. The quoted remarks ("marginal", etc.) are Hughes's judgments on whether the mean correlation values are significantly different or not. They reflect his experience in routinely evaluating the quality of the numerical guidance for the five-day forecast program. In view of the rather large case-to-case variability of the scores any conclusion based on this probably too small sample would be suspect about the best that can be said is that neither model shows any systematic advantage over the other.

The precipitation scores, Table 6, tell a different story - as can be seen the Spectral model was consistently better and Hughes considered the Spectral forecasts a "big improvement" over the 7L PE. It is, to say the least, not clear why the 5-day precipitation scores should give such a pronounced advantage to the Spectral model while the day-to-day verification of Figures 21 and 22 suggested that the 7L PE was the better model. Let us accept the horse as a gift and behave accordingly.

MEDIUM RANGE FORECAST VERIFICATION

(Fran Hughes)

	Day 3 SLP Correlation				D+3 DN 500mb Correlation			
	N.A.		U. S.		N.A.		U. S.	
Date	SP	7L	SP	7L	SP	7L	SP	7L
16 Jan	52	50	20	-5	65	65	42	10
11 Apr	63	65	23	20	76	72	62	57
30 Aug	37	49	21	17	56	54	56	38
5 Sep	61	59	60	62	32	86	42	86
30 Sep	59	60	51	82	82	80	83	83
Mean	54.4	56.6	35.0	35.2	62.2	71.4	37.0	34.8
	"marginal"		"not signif."		"significant" (one case)		"not signif."	

Table 5

DAY THREE PRECIPITATION SKILL SCORES

Date Made	Spectral	7LPE
Jan. 16, 1979	9	-1
Apr. 11, 1979	16	-22
Aug. 30, 1979	14	11
Sep. 05, 1979	60	47
Sep. 30, 1979	26	23
Average	25.0*	11.6

5-DAY MEAN PRECIPITATION SKILL SCORES

Date Made	Spectral	7LPE
Jan. 16, 1979	-6	-2
Apr. 11, 1979	29	10
Aug. 30, 1979	14	10
Sep. 05, 1979	33	25
Sep. 30, 1979	39	16
Average	21.8*	11.8

"Big Improvement" - Hughes

Table 6

iv. MOS Forecasts

The last bloc of objective evaluations were undertaken by the Computer Systems Branch of the Techniques Development Laboratory under the direction of Paul Dallavalle. The complete set of forecasts from both models was made available to the TDL, and a rather large aglomeration of Model Output Statisitics (MOS) programs (the ones usually run from the 7L PE output) were re-run from both forecasts models, and the results verified by TDL.

In order to facilitate running the MOS programs from the Spectral model portions of the output from the model were made to resemble 7L PE output (and were so identified). In particular the wind, potential temperature, relative humidity, and vertical motion (ω) from the lowest layer of the Spectral model were simply relabled as "Boundary Layer" quantities so as to fool the MOS programs. This was recognized to be a possible source of errors - the 7L PE Boundary Layer is 50 mb thick, the lowest Spectral model layer is 83 mb deep. Since the reidentification of the layer quantities took place without any adjustment to their values, the temperature, for example, might be a degree or so warmer than it "should" be. This depends, however, on what assumptions you chose to make about the temperature lapse (or the wind shear). There seems little guidance available for such choices, and we opted for the "choice" of doing nothing.

Once the MOS forecasts were run, TDL also undertook the compar - tive evaluation - here is a compilation of TDL's conclusions relating to the results:

- Trajectory Model: at 1000 mb the 7L PE was better but only by 0.4°C root-mean-square error (RMSE); at 850 & 700 mb - little difference.

- Bonneville Forecasts (max/min temperatures 60 to 96 hours): neither model could be judged better.

- Alaska Ceiling, Visibility, Cloud amount - 7L PE perhaps slightly better.

- Alaska Wind Speed and Direction - neither model better.

- Atlantic & Pacific Waves - both forecasts about the same - neither verified.

- Ocean Buoy Wind Speed and Direction - both models about the same.

- West Coast Station Wind Directions - models were different but neither were verified.

- Alaskan max/min Temperatures - at 24 hours the 7L PE was 1°F better in terms of mean error (bias), for 36-60 hours the models were essentially the same; in terms of mean absolute error the 7L PE was 0.5 to 1° F better for all projections.

- Alaskan Probability of Precipitation (POP) - 7L PE "considerably better" than Spectral at all three projections.

The last two items are of the greatest concern.

Table 7 exhibits TDL's verification scores of the Alaskan POP forecasts in terms of Brier Score. Small differences in Brier Scores are not insignificant and translate into respectable differences in percent improvement over climatology, differences indeed large enough to be of some concern.

Table 8 shows the 10 case climatology of the model's forecasts. These results confirm what was noted in the objective precipitation verifications in the lower 48 states - the Spectral model forecasts less occurances of precipitation than the 7L PE, whether measured by areal coverage or probability of occurance. In Alaska, as in the eastern U. S., this is not a good thing as the 7L PE is already underforecasting there.

Probability of Precipitation Brier Scores Observed for 14 Alaskan Stations

During the 10 Test Cases

Model	Forecast Period		
	1	2	3
Spectral	.149	.160	.205
7 LPE	.126	.130	.198

Table 7. TDL Verifications of Alaskan PoP Forecasts

A Comparison of the Average PoP Forecasts With the Observed Relative Frequency

Model	Average First Period PoP Forecasts	Observed Relative Frequency of Precipitation
Spectral	.25	.38
7L PE	.30	

Table 8. "Climatology" of Alaskan Precipitation Forecasts, TDL

D. Evaluation Summary - Where the Problems Are.

There appear to be five main (and not unrelated) areas of concern where the 7L PE showed some superiority over the Spectral model:

- Upper tropospheric winds, in particular the wind speeds, are systematically underforecast in the Spectral model. (They are underforecast in the 7L PE model too, but not so badly).
- Upper tropospheric temperatures show a positive (warm) bias in the Spectral model, again more so than the 7L PE.
- The stratospheric forecasts are not particularly good in either model and there is a suggestion that the Spectral is slightly the worse.
- The quantitative precipitation forecasts show a clear preference for the 7L PE model, with the rather puzzling exception of the medium range ("Day 3") specialized products.
- The Alaskan region MOS forecasts, in particular the max/min temperatures and PoP, show a preference for the 7L PE model.

These various impediments to implementation did not, of course, remain unrecognized until the full set of the 10 forecasts were completed. As the problems came to light, remedial activities were undertaken, aimed at eliminating the problems: a number of possibilities were explored. The results of these explorations constitute the content of the next section.

V. Modifications to the Spectral Model

The first three areas of concern all, of course, suggested that there was something lacking in the vicinity of the tropopause - in simplest terms the 7L PE has a tropopause and the Spectral does not. Related to this was the question of the vertical resolution of the models in the vicinity of the tropopause and above. Figure 23 shows, on the left, the vertical layer structure of the 12 even layer version of the Spectral model (with a logarithmic pressure coordinate); the two central columns are typical 7L PE structures for high level (tropical) and temperate zone tropopause values. The topmost Spectral layer reaches to infinity (zero pressure) while the 7L PE ends at 50 mb.

In the immediate vicinity of the tropopause it would appear that the Spectral has as good or better resolution than the temperate 7L PE; however the 7L PE has "engineering" features built into the output sections that force a temperature minimum and wind speed maximum at the tropopause. This min/max is reflected in the temperatures or winds interpolated to nearby mandatory levels. Some experimentation a few years back (unreported upon) suggested that 14 even layers (without a tropopause) achieved sufficient resolution to compensate for the loss of the tropopause-engineered maximum and minimum of winds and temperatures. Some of the decrease in forecast quality in the winds and temperatures at 250 mb was attributed to the loss of the engineered tropopause in the Spectral model.

The right hand side of Figure 23 shows what was proposed for the new vertical layering of the Spectral model - six 50 mb layers from 300 mb on up to zero pressure and a quasi-uniform thickness variation below. In the vicinity of the tropopause this is equivalent to a 20 (even) layer structure, presumably, one hopes, enough. We experienced no qualms in reducing the resolution in the lower troposphere 33% (6 even layers below 500 mb to 4 rather unequal ones) because of the general similarity of the Spectral and 7L PE in the lower levels - if 6 layers showed no improvement over the two tropospheric layers plus the boundary layer of the 7L PE, cutting the six to four should make no difference. This also gave us the opportunity to reduce the lowest layer thickness from 83 to 75 mb, bringing it closer in line with the boundary layer of the 7L PE.

In the stratosphere further benefits are anticipated - the resolution, above 250 mb, has gone from three to five layers (or four useful layers), somewhat better than the 7L PE. Earlier experience with other NMC models had indicated that three stratospheric layers are marginally sufficient (only two allow for excessive noise generation); we hope four will be better still.

Resolution was not the only suspected culprit, particularly with reference to the winds. Some calculations by both SEB (and independently by TDL) showed a steady decrease of kinetic energy throughout the model at various levels during the forecasts; simple histogram counts of the number of forecast gridpoints with wind speeds in 5 mps bands also showed a clear reduction in the number of strong wind points as the forecast progressed. Two sources of the presumed excessive smoothing were considered, and experimented with. The results indicated that the linear (∇^4)

smoothing coefficient should be reduced by an order of magnitude and that the time stepping be changed from "backward implicit" to "centered implicit". The latter step was coupled with the introduction of non-linear normal mode initialization, which had been studied independently of the present evaluation tests and found to be of considerable value for the short range (6-12 hour) forecasts (Cressman, 1980).

A potential bonus from the rearrangement of the layers was realized in the large scale precipitation forecast section of the model. In that section there is a parameterization of the saturation criterion for each moist layer which is a function of the particular layer and the lowest layer temperature. The scheme was originally developed for NMC's (now defunct) 9 Layer Global Forecast Model (Stackpole, 1978) which had 6 even tropospheric layers, of which the lowest 5 contained moisture. The new Spectral layering is considerably closer to the 9 layer model structuring than was the old; thus the parameterization, which was carried over with minimal alterations, may be better tuned to the model than before. At least that is the hope.

While considering these possibilities an error of misunderstanding was unearthed. It is necessary, to avoid initial latent heat excesses, to somehow limit the initial relative humidity in the σ-layers to be not greater than the layer-by-layer saturation criteria. The saturation criteria is generally less than 100%. We discovered that this limitation had been achieved by simply truncating the humidity at 80% in all the layers. This is too heavy handed. What was intended and corrected in the new Spectral was that the humidity be scaled (by multiplication by the fractional saturation criterion) in each layer using the criterion appropriate to that layer. For instance the saturation criterion in the lowest layer is 100%. The old, erroneous, system would cause the initial humidity there to be less than saturation even if the observed humidity were truely at saturation. Thus the model would start out too dry. The new system corrects this defect.

Once all these changes and corrections were in place and tested, the complete set of 10 cases were rerun with the new Spectral model, the subjective jurors were invited to assess the new version (although not in as great detail as the first time around); the various objective verifications and the MOS codes were run and evaluated. The next section describes the results.

VI Results - Second Round

A. Subjective Evaluations

For the second round of subjective evaluations a considerably simpler questionnaire (See Appendix II) was given to the original evaluators. They were merely asked to compare the old and new Spectral model forecasts, with reference to the 7L PE and verifications, of course, and check one of three choices:

- "New Spectral is better and has overcome my previous objections"
- "Little Different"
- "New Spectral is worse."

These are not very penitrating examinations of the opinions of the evaluators but, in the light of the more detailed examinations in the first round, they were felt to be sufficient. Again the results were tabulated and grouped by field of specialization.

i. Aviation forecasting

R. McCarter checked "new Spectral better" in all ten cases but very carefully crossed out the phrase "and has overcome my previous objections" because, he said, he hadn't had any. In conversation he did mention that the jets were forecast with higher speeds in the new Spectral model and that was in the right direction, he said.

ii. Marine forecasting

H. Brown selected "Little different" 4 times, "Spectral better..." 6 times, but went to the trouble of inserting "little" before "better" for three of those six cases. Evidently he didn't find much to become excited about; but little change in the lower troposphere was what we anticipated anyway.

iii. Quantitative Precipitation forecasting

D. Olson opted for one "better", seven "little different" and two "worse". The two "worse" cases were both poor precipitation forecasts in the first place and became somewhat more poor. The others were not particularly useful guidance, either, in both the old and new Spectral runs.

This is rather disappointing result and will serve as a goad for the modelers to give more attention to the precipitation sections of the model.

iv. General Quality & Credability

H. Saylor marked all 10 cases "little different" but remarked in conversation that he, too, noted, and was glad to see, the increased wind speeds in the jet streams at 250 mb.

v. Tropical & Southern Hemisphere

E. Carlstead reviewed both his previous comments and the new & old Spectral forecasts. His observations boil down to two points which were made repeatedly case by case:

a. The 1000 mb charts are essentially unchanged.

b. The jet stream winds are faster.

Since the evaluator's original comments were generally favorable, these two are a welcome addition to the pantheon.

vi. A Bonus: Two New Evaluators

As the second round of testing got underway G. P. Cressman stepped forward and volunteered to review all three forecasts (two Spectrals and the 7L PE) for the ten cases. In addition Prof. Lance Bosart of the Meteorology Dept. of the SUNY/Albany got wind of the tests and asked if he too could participate. We promptly sent him the maps for the new Spectral (only), the 7L PE, and the verifications. The Cressman evaluations came back in the form of remarks about each cases; Bosart used the original multi-category questionnaire (Appendix I) and offered comments as well.

Cressman gave each case a fairly detailed consideration, more than can be usefully reproduced here. However certain common themes are discernable in his remarks:

- At the surface the new Spectral is either equal to or better than both the old and the 7L PE.

- At 500 mb they are more generally equal, with the new Spectral occasionally better. The new Spectral appears somewhat more noisy than the old but it wasn't clear if this was a bad thing - the "noise" may very well be a better representation of true variability.

- At 250 mb the new Spectral is quite consistently better than both the old and the 7L PE.

- At 100 mb, all three are pretty much tied.

- The new Spectral is a clear winner for the tropopause pressure and wind shear.

- The precipitation forecasts are a very mixed bag: no clear preference emerges-one model will be better at one forecast hour, another model at another hour of the same case, etc. Frequently they are all bad.

It appears that Cressman's independent evaluation is in quite general agreement with the judgements of the other evaluators.

Since Prof. Bosart used the questionnaire form we can tabulate his results in the same manner as the others:

Aviation Forecasting (Bosart)

Category	Preference Tally		
250 mb Winds	NEW SPEC	7L PE	Ties
Atlantic	8	2	0
N. America	6	1	3
Pacific	7	0	3
Sub-Total	21	3	6
250 mb hts. & temps			
Atlantic	6	0	4
N. America	8	0	2
Pacific	5	2	3
Sub-Total	19	2	9
Grand Total	40	5	15

Table 9

(Prof. Bosart distained to consider the tropopause and stratospheric levels.) Clearly the results here show a clear preference for the (new) Spectral over the 7L PE, in most satisfactory agreement with the other objective and subjective evaluations.

Marine Forecasting (Bosart)

Category	24 HR NEW SPEC	24 HR 7L PE	24 HR Ties	48 HR NEW SPEC	48 HR 7L PE	48 HR Ties	24 & 48 HR Combined NEW SPEC	24 & 48 HR Combined 7L PE
SLP & Thickness								
Atlantic	4	4	1	8	2	0	12	6
Pacific	5.5	2	2	5	3	2	10.5	5
Subtotal	9.5	6	3	13	5	2	22.5	11
500 mb Ht & Vorticity								
Atlantic	3	3	3	4	3	3	7	6
Pacific	6	1	2	6	2	2	12	3
Subtotal	9	4	5	10	5	5	19	9
Grand Total	18.5	10	8	23	10	7	41.5	20

Table 10

Here the results differ somewhat from the in-house evaluation - the latter concluded that the two models were essentially of equal value; Prof. Bosart favors the Spectral model. In a letter that accompanied his returned questionnaire Bosart said:

> "The models are similar but the Spectral earns extra brownie points here and there for little circulation details that mean more to fussy (anachronistic?) synopticians like me."

Presumably the "little circulation details" did not mean as much to the more operationally oriented meteorologist (H. Brown) at NMC.

Quantitative Precipitation Forecasting (Bosart)

Category	24 HR			48 HR			Combined	
	NEW SPEC	7L PE	Ties	NEW SPEC	7L PE	Ties	NEW SPEC	7L PE
Rain/No Rain Coverage								
East	1	7	1	2	7.5	0	3	14.5
West	1	8	0	1	5	3	2	13
Subtotal	2	15	1	3	12.5	3	5	27.5
Quantitative								
East	1.5	8	1	2	6.5	2	3.5	14.5
West	1	7	3	1	6	3	2	13
Subtotal	2.5	15	4	3	12.5	5	5.5	27.5
Total	4.5	30	5	6	25	8	10.5	55

Table 11

(Prof. Bosart gave scant attention to the 84 HR forecasts)

The now all too familiar conclusion on the relative merits of the precipitation forecast sections of the models is once again apparent. Bosart recognized that the precipitation forecasts from the 7L PE are not critical to NMC's national forecast responsibilities and did not view the difficulties as a bar to implementation.

B. Objective Verifications

i. Heights, Winds, and Temperatures

All of the objective verification calculations were repeated for the modified (or new) Spectral model. There is little point in presenting the diagrams for all the various statistics, areas, levels, etc. to give a one-for-one comparison of the new and old model as many of the conclusions would be redundant. However we do show here a subset concentrating on the critical problem areas.

Figure 24 and Figure 25 are actually not problem area verifications. They compare to Figure 1 and Figure 2 and show that the 500 mb verifications have scarsely changed at all. At least we have done no harm.

In Figures 24 and 25 (and the others to follow) the ordinate has been marked as "OPNL" (for "operational"), a bit presumptively perhaps, but it serves to distinguish the new from the old, and to give a hint of the concluding recommendations of this essay.

Figure 26 and Figure 27 are the verifications for the first of the problem areas, the 250 mb temperatures (cf Figures 5 and 6). Figure 26 shows the temperature bias has not been entirely eliminated, unfortunately, but is now on the order of only half a degree or so. Figure 27 is much more pleasing showing that the previous preference for the 7L PE in terms of RMS temperature error has been eliminated.

Figures 28 and 29 (cf Figures 7 and 8) are even more pleasing, showing as they do that the 250 mb mean speed error comparison now favors the Spectral model without any alteration to the tie scores of the RMS vector wind errors.

The previous 6 diagrams were for the Northern Hemisphere - the remaining ones use the North American verification network. The first two of those, Figures 30 and 31, are for the temperatures at 250 mb; the previous bias (Figure 14) and RMS temperature errors (Figure 13) that appeared to favor the 7L PE (slightly) now are either neutral or (slightly) favor the Spectral model.

Figure 32 (cf. Figure 16) is also pleasing showing that the new (OPNL) Spectral has most definitely bested the 7L PE in the wind speed bias category. This serves as a confirmation of the hemisphere results.

A special note about the wind verifications: recall that the 7L PE model as run in these tests had less diffusion than the operational 7L PE. This caused the winds in the test 7L PE to be somewhat stronger (and therefore better in terms of speed bias) than the operational 7L PE. That the Spectral winds are (now) better still is a very positive sign for the model.

When we turn to 100 mb the hope for improvements of the new Spectral over the old are realized, but not so dramatically as at the lower level. Figures 33 and 34 (cf. Figures 17 and 18) exhibit the mean and rms temperature errors. They are still quite case-by-case variable but there are certainly indications that the new Spectral shows some improvement over the old and is now "equal" to the 7L PE.

The same conclusion is true for the wind verifications, Figures 35 and 36 (cf. Figures 19 and 20). Note particularly in the speed verification that the original verdict of Spectral speeds too fast has been somewhat ameliorated in the new model. Evidently the increased resolution has, in some way, properly isolated the 100 mb/lower stratosphere regions from the strong jet stream below, at least partially.

Some tropical verifications of 100 mb, not shown here, where the tropopause is in the vicinity of 100 mb, do not show the new resolution to any advantage. The 7L PE and <u>both</u> Spectrals show, for example, a 4 to 5°C warm bias for the 100 mb temperatures. There is just not sufficient resolution (even with 50 mb layers) to do justice to the tropical tropopause and stratosphere. Improvements here will probably have to await the availability of larger computers.

ii. Precipitation

R. Hirano undertook the objective precipitation verifications as before. Figure 37 and 38 are copies of Figures 21 and 22 with the "OPNL SPECTRAL" results added, denoted by open squares. It only takes a moment's perusal of the figures, paying particular attention to the relative positions of the dots and boxes, to see that not much has changed - the objective verifications agree in that respect with the subjective evaluations. About the only glimmer of hope is seen in the lowest graphs of Figure 38: the excessive amounts of precipitation (in the still too small areas) are slightly, and consistently, reduced. At least that's something.

iii. Medium Range Forecasts

The statistical evaluations of the five medium range forecasts (to 6 days) were again undertaken by F. Hughes. Tables 12 and 13 show the old Spectral scores (reproduced from tables 5 & 6) and the new, side by side. As can be seen the new Spectral scores contain no surprises, nor any disappointments. The new ones are as case by case variable as the old (and probably about as meaningful); the slight increase in the precipitation skill scores (on the average) is heartening but probably not meaningful.

MEDIUM RANGE FORECAST VERIFICATION

(Fran Hughes)

	Day 3 SLP Correlation				D+3 DN 500mb Correlation			
	N.A.		U. S.		N.A.		U. S.	
Date	OLD SP	NEW SP	OLD SP	NEW SP	OLD SP	NEW SP	OLD SP	NEW SP
16 Jan	52	54	20	12	65	72	42	38
11 Apr	63	59	23	14	76	77	62	63
30 Aug	37	31	21	10	56	49	56	43
5 Sep	61	66	60	63	32	54	42	90
30 Sep	59	61	51	54	82	82	83	88
Mean	54.4	54.2	35.0	30.6	62.2	66.8	37.0	64.4

Table 12

DAY THREE PRECIPITATION SKILL SCORES

Date Made	Old Spectral	New Spectral
Jan. 16, 1979	9	5
Apr. 11, 1979	16	16
Aug. 30, 1979	14	11
Sep. 05, 1979	60	60
Sep. 30, 1979	26	36
Average	25.0	25.6

5-DAY MEAN PRECIPITATION SKILL SCORES

Date Made	Old Spectral	New Spectral
Jan. 16, 1979	-6	7
Apr. 11, 1979	29	24
Aug. 30, 1979	14	17
Sep. 05, 1979	33	29
Sep. 30, 1979	39	42
Average	21.8	23.8

Table 13

iv. MOS Forecasts

Last but not least portions of the MOS forecasts, those dealing with the Alaskan forecasts, were rerun. Because of a number of difficulties it was possible to verify only the max/min temperatures and the PoP forecasts. The bias error of the max/min was unchanged for the 24 and 36 hour forecasts and became about 1° F colder than the 7L PE at 48 and 60 hours. The RMS temperature error was essentially unchanged with the new Spectral model.

Brier Score
New Spectral Model
14 Alaskan Stations

Period	1	2	3
Score	0.134	0.150	0.196

Table 14

The PoP Brier scores are in Table 14. Comparison with Table 7 shows a good improvement at all three periods (the Brier score is an error score, so the lower the better) even to the extent of surpassing the 7L PE score at the third (36-48 hour) period, not, unfortunately, by much.

C. Summary and Recommendation

Of the five main areas of concern (Section IV-D, above) it appears that the most critical ones have been overcome and the level of concern greatly reduced for the others:

- The Spectral now forecasts the upper tropospheric winds <u>better</u> than the 7L PE.
- The upper troposphere temperature bias has been reduced to within half a degree or so of the 7L PE bias.
- In the stratosphere the Spectral model now either has a (very slight) edge over the 7L PE or the two are equal in quality. Neither are terrifically good.
- The relative quality of the precipitation forecasts remain the same.

- The Alaskan MOS forecasts still show a slight preference for the 7L PE but improvements in the PoP are apparent.

In the tropics and southern hemisphere there is good evidence that the Spectral model is capable of making useful forecasts.

The recommendation is, of course, to implement the Spectral model in NMC's operations as the replacement (after more than 14 years) for the Shuman-Hovermale 7L PE model.

Upon (and prior to) such implementation we shall give particular R+D attention to two aspects of the model: there will be efforts to tune the precipitation sections, to find what is lacking and make repairs, and radiation calculations of an as yet undetermined level of sophistication will be introduced. As with all new models there will no doubt be a period of rapid change and development as further problems or opportunities for improvement come to light.

VII. Coda - The Cases and Their Selection

As mentioned 'way back in Section III the ten selected cases were chosen partially on the basis of their availability in the NMC archive - the cases got to the archive usually because of some special event. I thought it worth while to look at the particular special event of each case (listed in III-B) to see what the Spectral may have done for us.

1. The New England Blizzard of 78

By 12Z Monday 6 Feb 78 (48 hours into the forecast) the verification showed a nice strong looking 1003 mb low some 2° east of Cape Hatteras. The storm subsequently moved north and closed New England for a week. The 7L PE grew a rather bland 1012 low some 5° east of St. Augustine while the Spectral placed a 1009 low 1° east of Hatteras (and a companion 1010 low 4° east of Miami). The Spectral wasn't perfect but was a clear improvement over the 7L PE.

2. Atlantic Blocking

The development of the block that was of interest occured some 5 days into the forecast - at 48 hours both models shared the same good and bad qualities unrelated to the future block.

A series of tests with the Spectral model, not immediately related to the question of implementation, showed that the successful forecast of the block depended as much on the initial analysis as on some of the physical/ numerical details of the forecast model. This will be reported upon elsewhere.

3. Lee of the Rockies Overdevelopment

The overdevelopment appeared in the 48 hour 7L PE forecast as a 994 mb low centered on the Wyoming-Montana border, with a 998 mb trough reaching to Illinois, while the verification showed only a 1001 mb low there with a trough reaching to Kansas. The Spectral model also produced a 994 mb low (in N. W. South Dakota) - certainly no improvement over the 7L PE.

4. Cross Contour Flow

The cross contour flow that caught the attention of one of the selectors occured in the first 12 hours of the forecast and didn't appear later. There was however some mid-Pacific cross contour flow - not particularly extreme - that both models exhibited to about an equal extent.

5. Precipitation Forecast Problems

As could be anticipated from earlier remarks, the Spectral model did not do anything to improve the poor rain forecasts.

6. Locked in Low - Initial Time 12Z 9 May 79

The series of analyses show, at 12Z 9 May, a 997 mb low at the extreme end of the Oklahoma panhandle which by 12Z 10 May has filled somewhat, to 1005 mb, and moved to Childress, Texas. There is also a warm frontal trough reaching to Wisconsin. Twenty four hours later the warm front trough has become the main low with a center of 1002 mb just north of Grand Portage, Minn. The 500 mb charts show a parallel sequence: starting with a 5450 m center low over southern Utah, to (24 hrs later) a trough, with no closed low, extending from East Montana to Arizona, to (24 hrs later) a closed low, 5430 m central value , over Winnipeg with a trough to Arizona. Both the 7L PE and Spectral did fine through 24 hours but failed miserably at 48 hr: they both left closed 500 mb lows over Arizona and no trace of the Winnipeg - Grand Portage system. Indeed they put ridges in that area. Clearly no triumph for either model.

7. 5 Day Forecast Case

I defer the evaluation of this to the medium range experts - judging from the scores in tables 5, 6, 12 and 13, the Spectral performed a little bit better than the 7L PE.

8. Hurricane David

A hurricane is a rather small scale feature for either numerical analysis or forecast systems to handle; both models placed low pressure areas (two weak and too broad scale, of course) about where they belonged and also produced satisfactorily generous rain amounts along the east coast. Neither model seems preferable.

9. Another Precipitation Case

As before the Spectral showed no improvement over a very poor 7L PE precipitation forecast. Perhaps the analysis is the root cause in this case as the pressure pattern forecasts were quite similar, and closer to each other than to the verifications.

10. The D. C. Blizzard of '79

This "Blizzard" was actually a small but intense storm that generated a lot of local excitement (and broken tree limbs) by dumping some 5 inches of wet snow on trees that were still in leaf. The Blue Ridge in Virginia was especially hard hit. This kind of storm is a challenge for any model - both models show reasonably satisfactory pressure pattern forecasts and generally inadequate precipitation forecasts. There is no particular advantage of either model over the other.

Acknowledgements

The performance of a compartive study such as this one involves many people; particular thanks go to:

Joe Sela, Ken Campana and Dennis Deaven for setting up and showing me how to run the forecast models; A. J. Desmarais for guidance in setting up a fail-safe forecast execution and archiving system; R. Hirano for maintaining the case library, and preparing the QPF maps; the members of NMC's Forecast Division, named in the text, plus G. Cressman and L. Bosart, for the subjective evaluations; Don Marks and the late John Horodeck, for setting up and running the objective verifications; P. Dalavalle and other TDL people for the MOS forecasts and verifications; C. Vlcek for energy statistics calculations; Carl Amorose for a vast amount of Varian map production; Phil Hovey for an even vaster amount of map duplication and distribution plus preparation of all the figures; and Dawn Starosta for all the typing. Whew!

FIGURES 1 - 38

FOLLOW

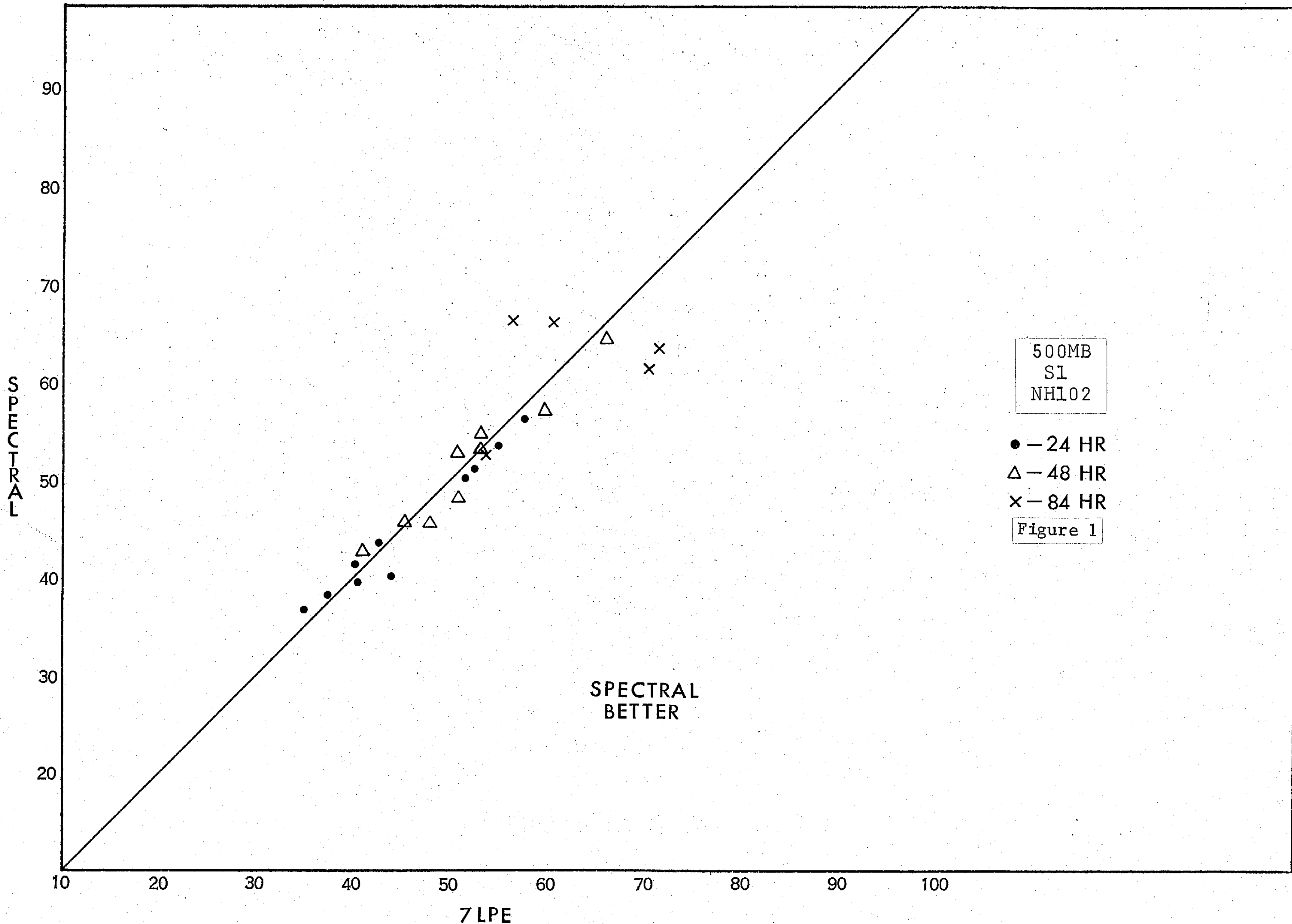
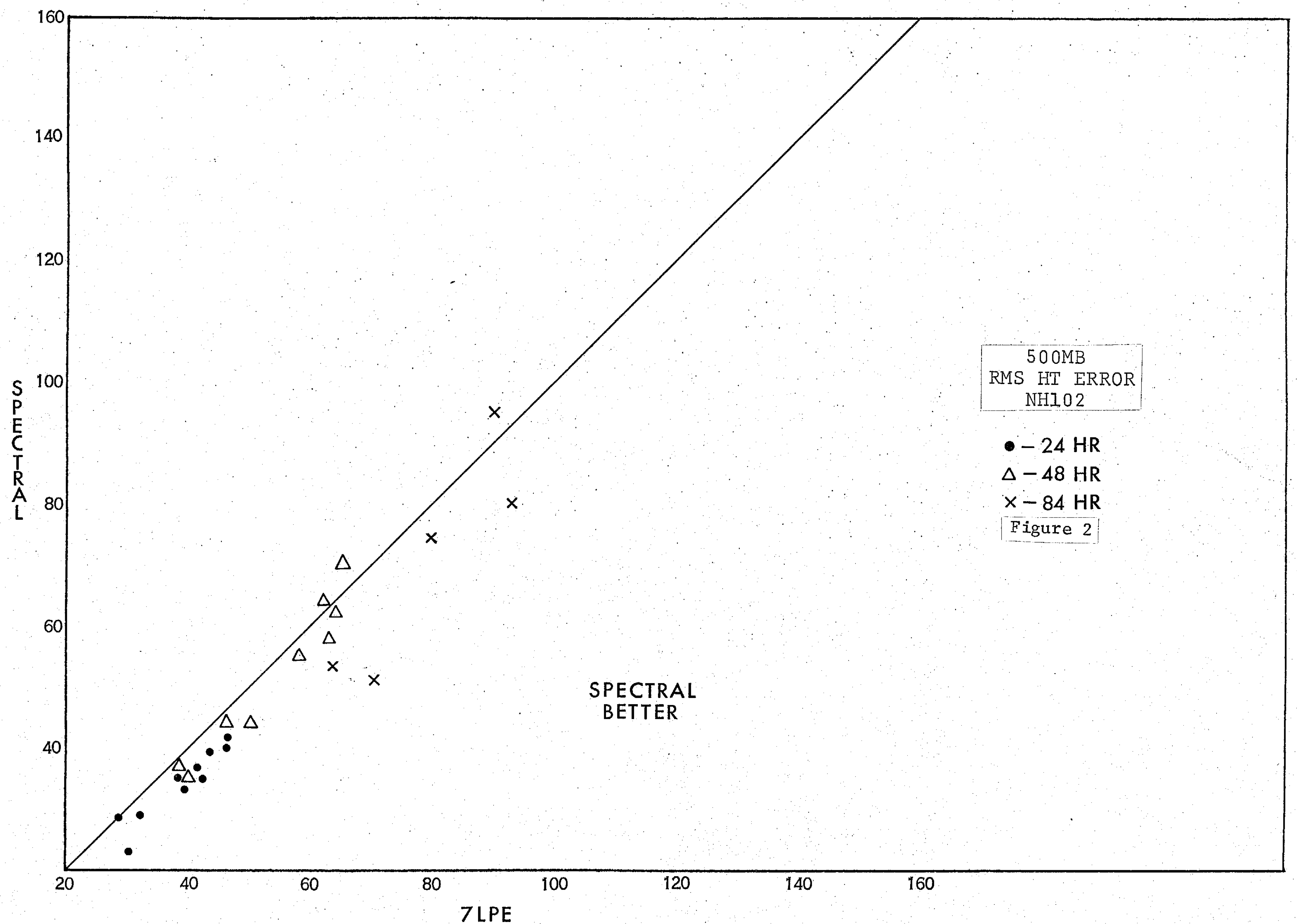
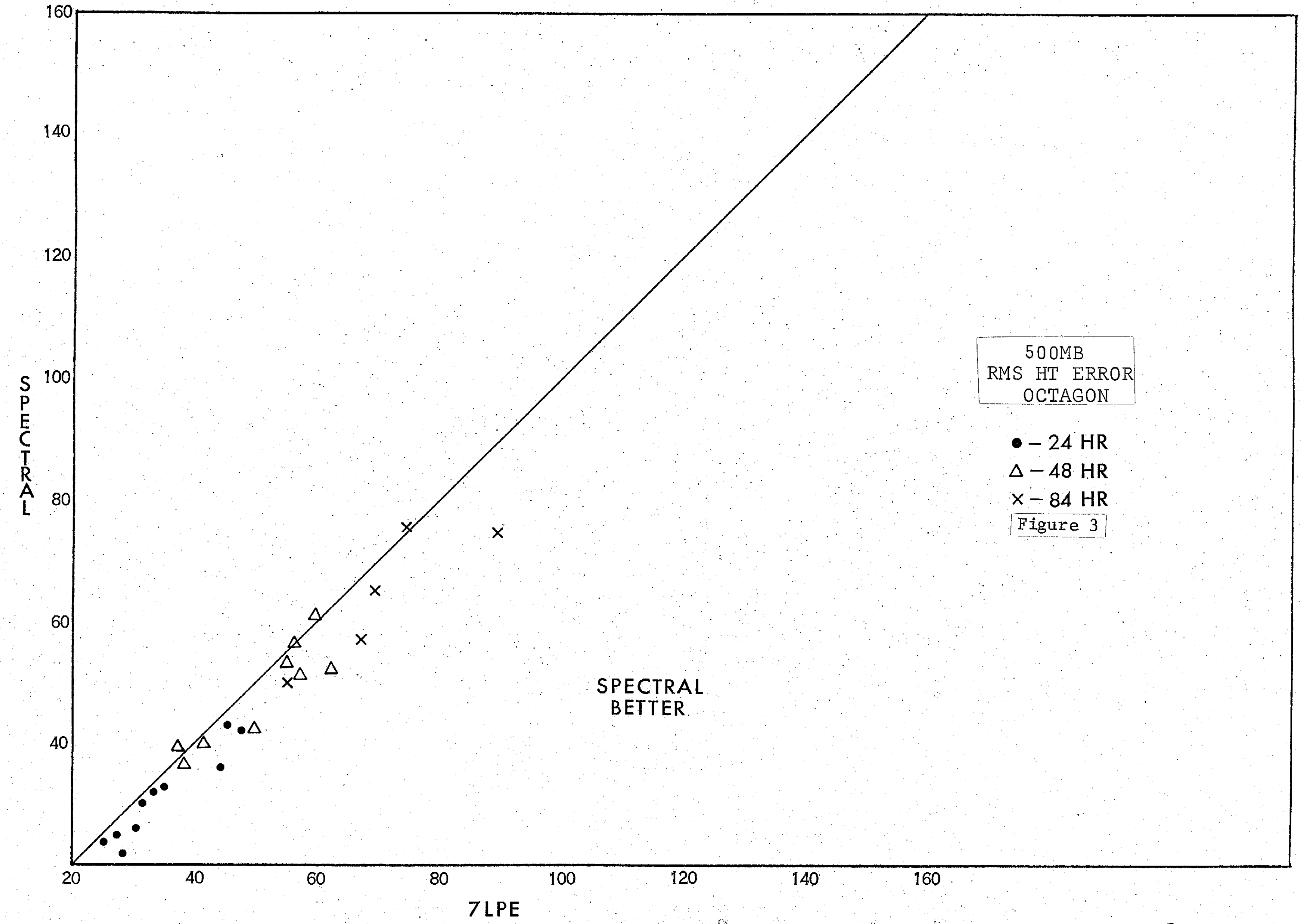

SPECTRAL
90
80
70
60
50
40
30
20
10
20
30
40
50
60
70
80
90
100
7 LPE
SPECTRAL BETTER
500MB
S1
NH102
● — 24 HR
△ — 48 HR
× — 84 HR

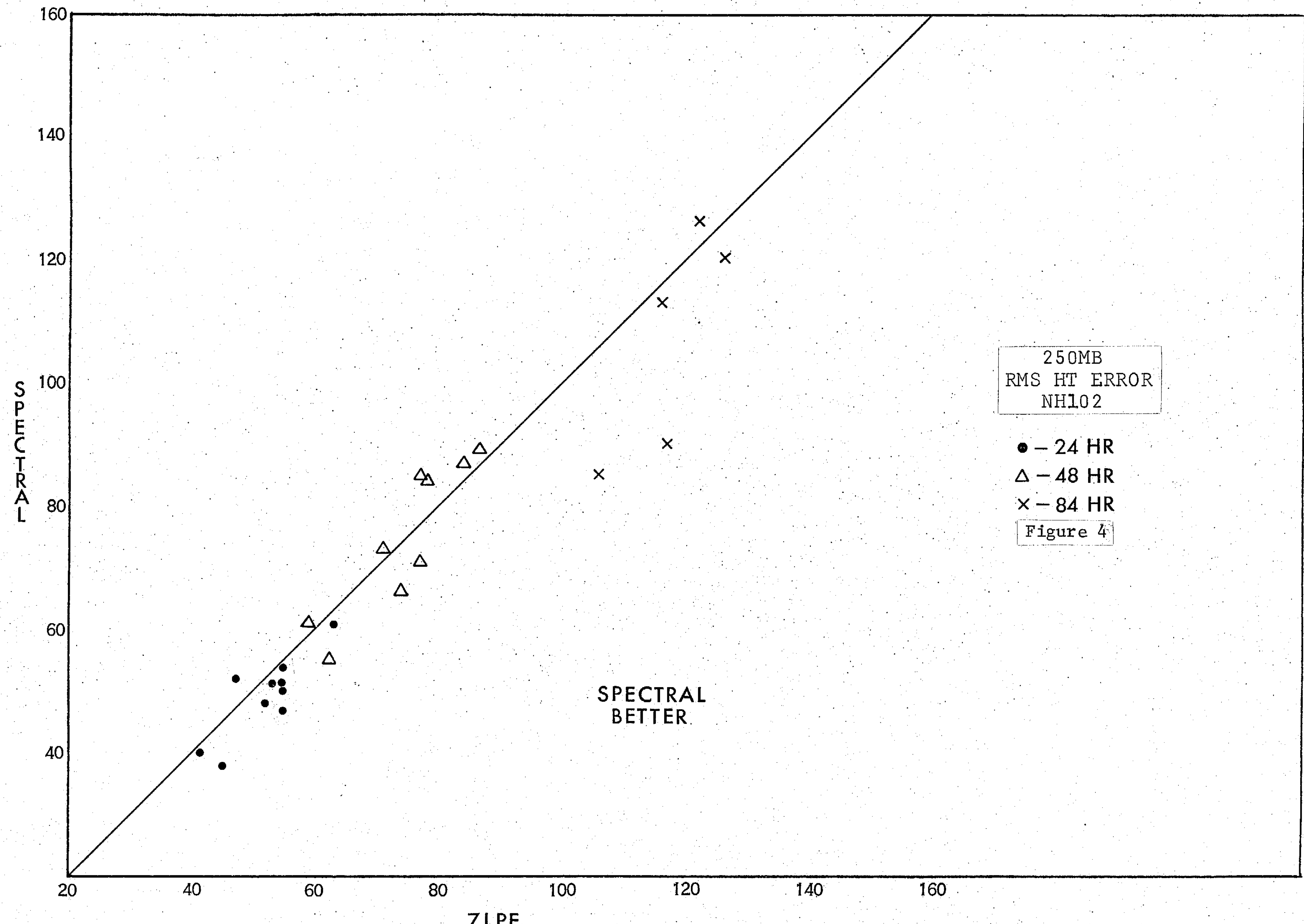
Figure 1

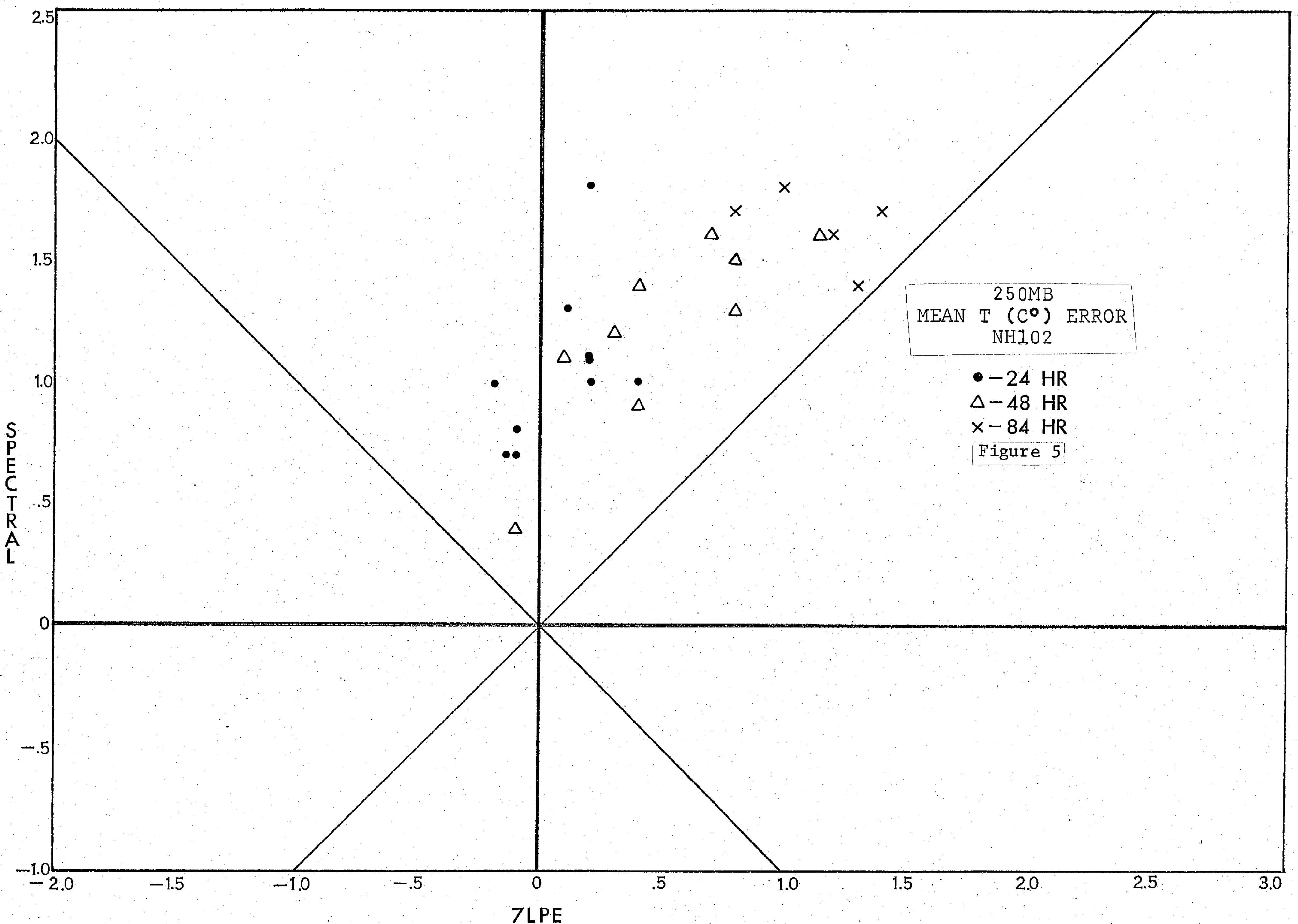

SPECTRAL
7LPE
500MB
RMS HT ERROR
NH102
● – 24 HR
△ – 48 HR
× – 84 HR
SPECTRAL BETTER
160
140
120
100
80
60
40
20
40
60
80
100
120
140
160
Figure 2



500MB
RMS HT ERROR
OCTAGON
● – 24 HR
△ – 48 HR
× – 84 HR
SPECTRAL BETTER
SPECTRAL
7LPE
160
140
120
100
80
60
40
20
40
60
80
100
120
140
160

Figure 3


160
140
120
100
80
60
40
SPECTRAL
20
40
60
80
100
120
140
160
7LPE
SPECTRAL BETTER
250MB
RMS HT ERROR
NH102
• – 24 HR
△ – 48 HR
× – 84 HR
Figure 4



250MB
MEAN T (C°) ERROR
NH102
• – 24 HR
△ – 48 HR
× – 84 HR
Figure 5
SPECTRAL
7LPE
2.5
2.0
1.5
1.0
.5
0
–.5
–1.0
–2.0
–1.5
–1.0
–.5
0
.5
1.0
1.5
2.0
2.5
3.0

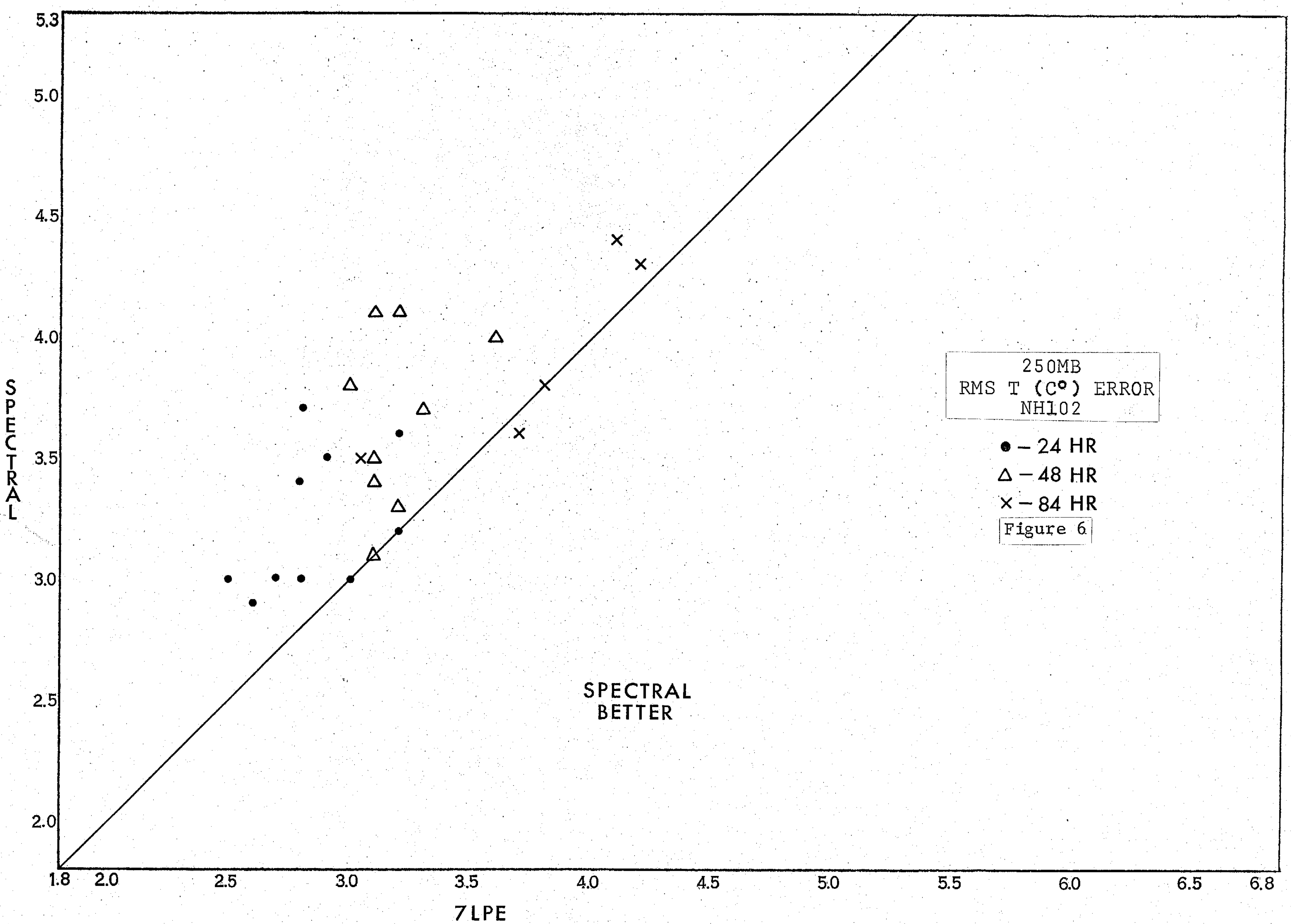
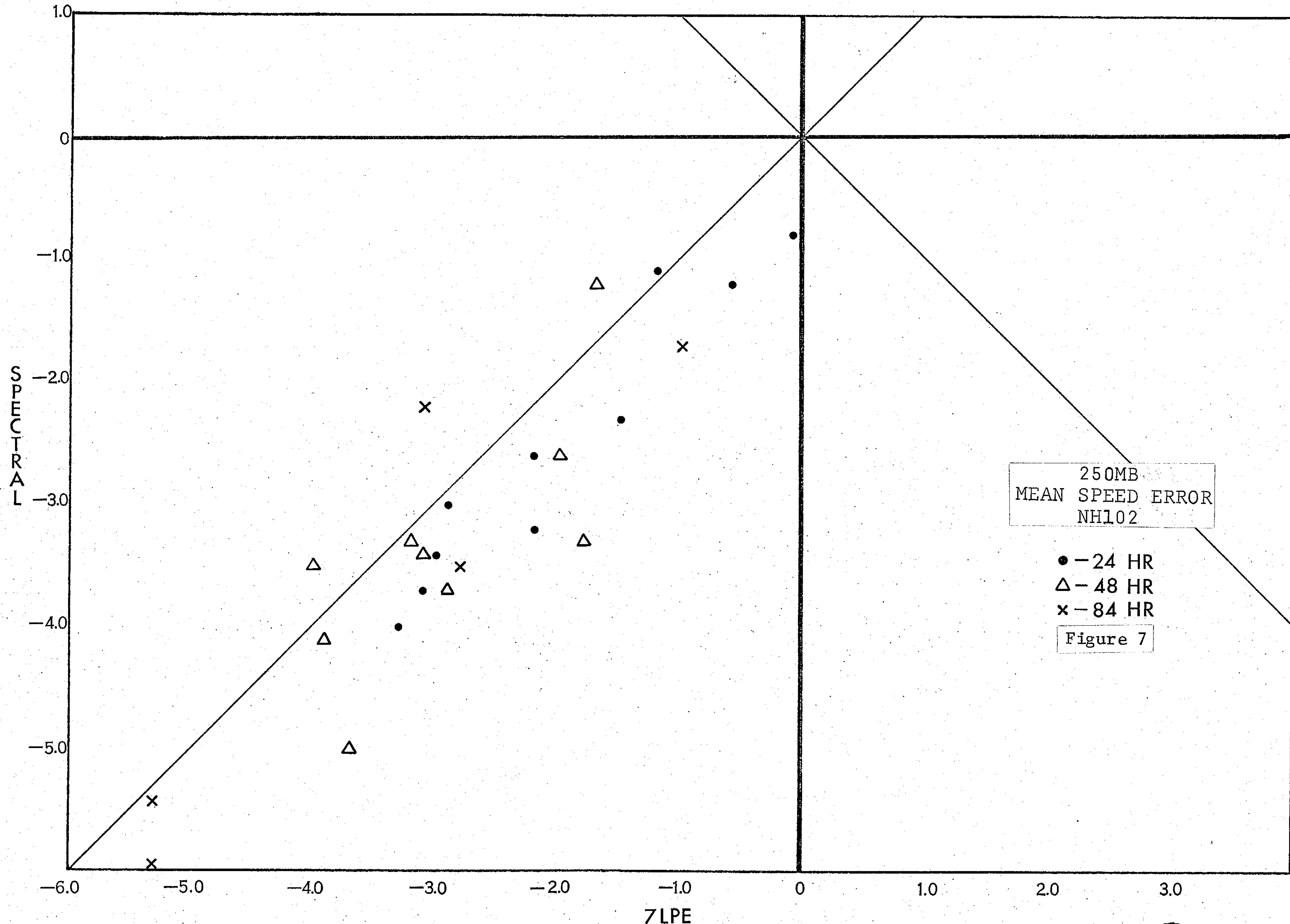
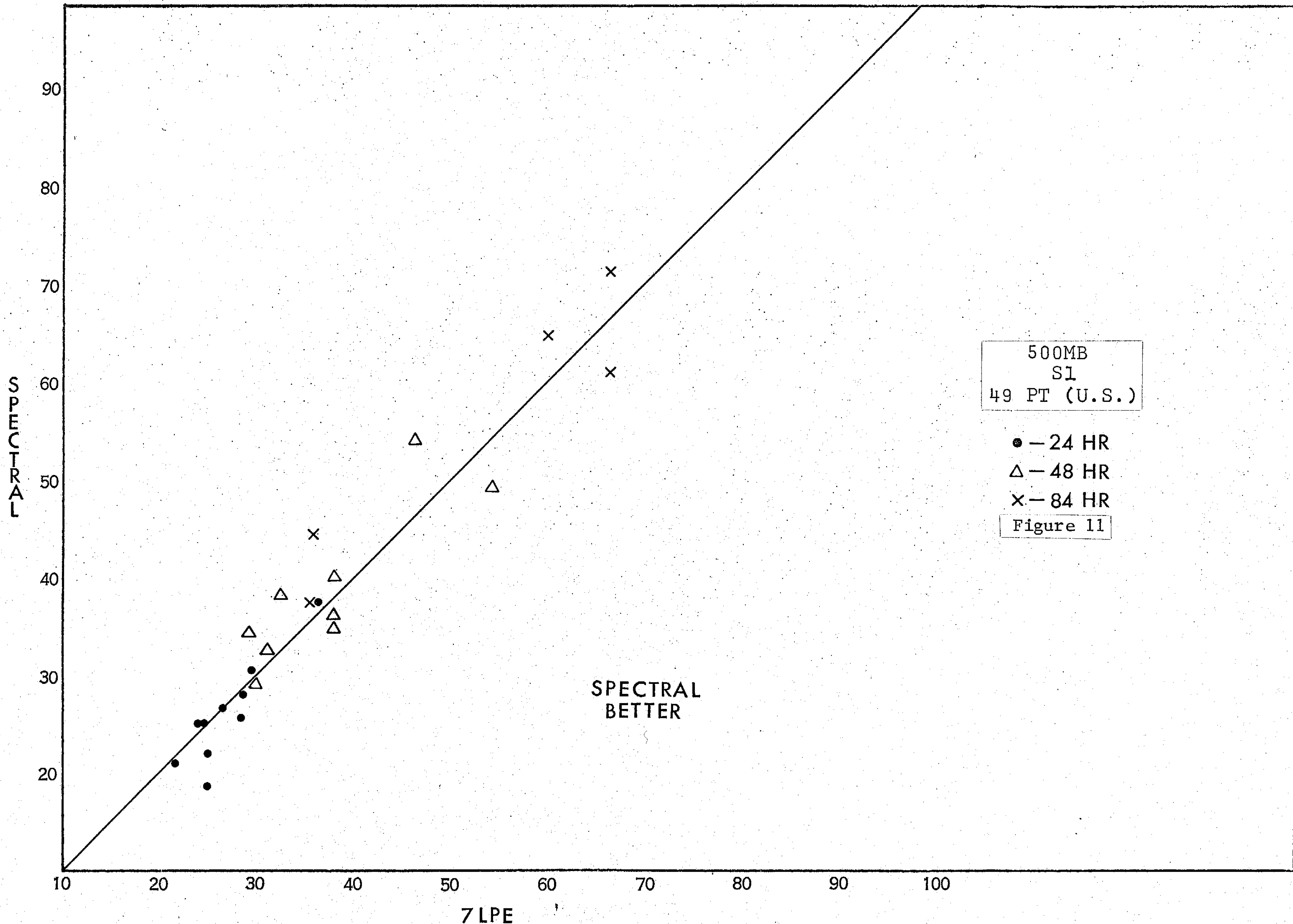

5.3
5.0
4.5
4.0
3.5
3.0
2.5
2.0
SPECTRAL
1.8
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
6.8
7LPE
250MB
RMS T (C°) ERROR
NH102
• – 24 HR
Δ – 48 HR
× – 84 HR
SPECTRAL
BETTER

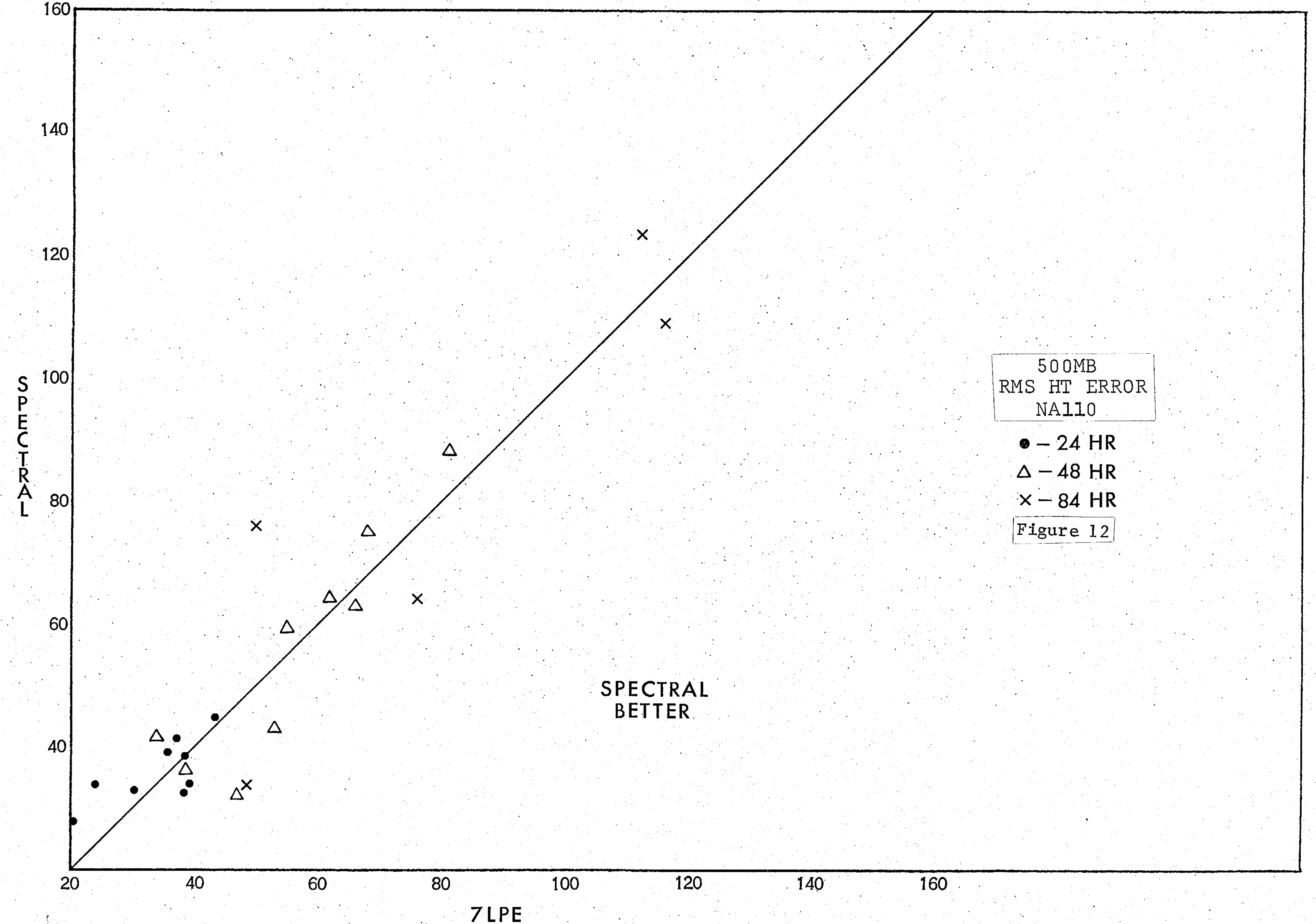
Figure 6



1.0
0
−1.0
−2.0
−3.0
−4.0
−5.0
SPECTRAL
−6.0
−5.0
−4.0
−3.0
−2.0
−1.0
0
1.0
2.0
3.0
7LPE
250MB
MEAN SPEED ERROR
NH102
● − 24 HR
△ − 48 HR
× − 84 HR

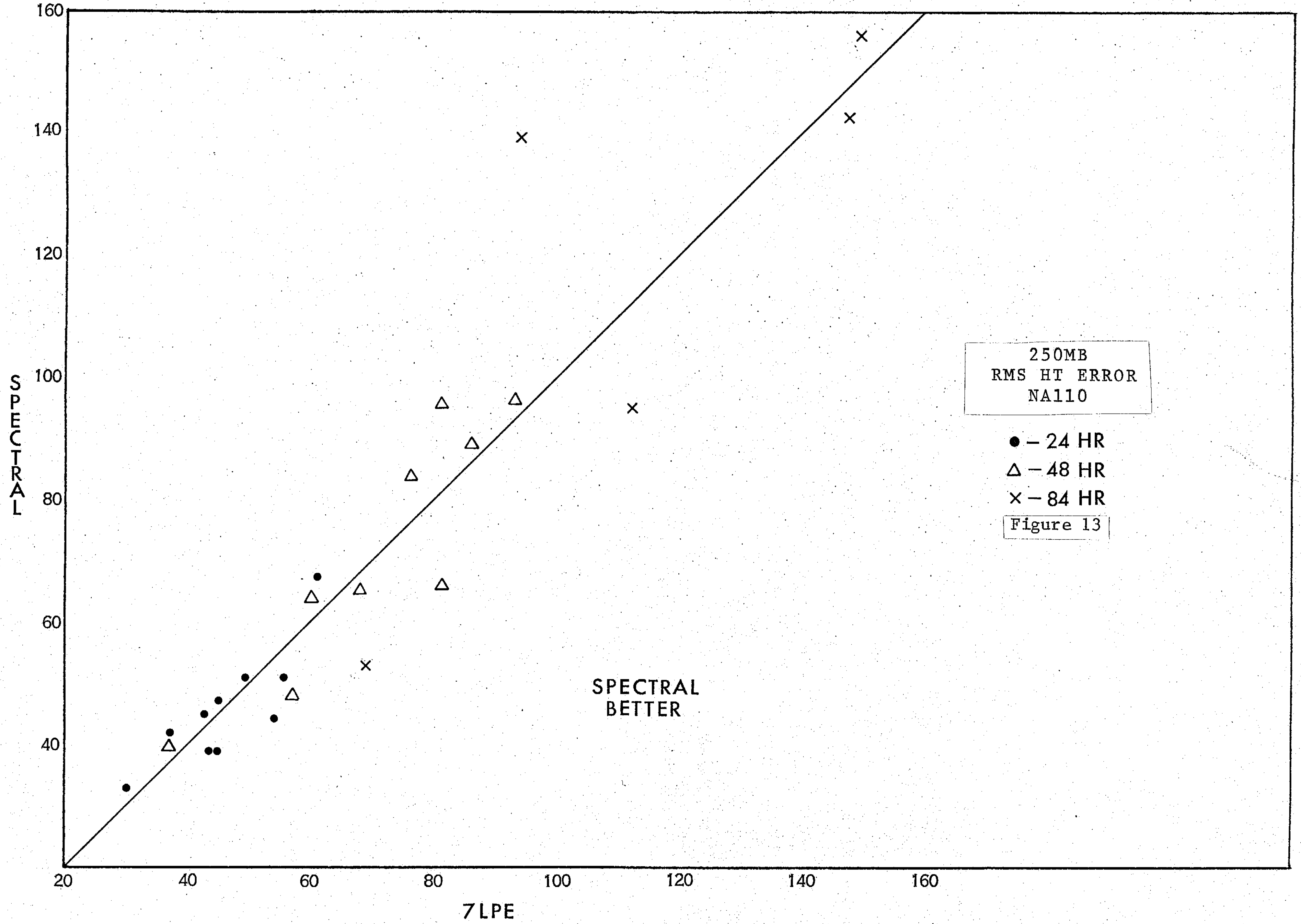
Figure 7


500MB
S1
49 PT (U.S.)
• — 24 HR
△ — 48 HR
× — 84 HR
SPECTRAL BETTER
SPECTRAL
7 LPE
10
20
30
40
50
60
70
80
90
100
Figure 11



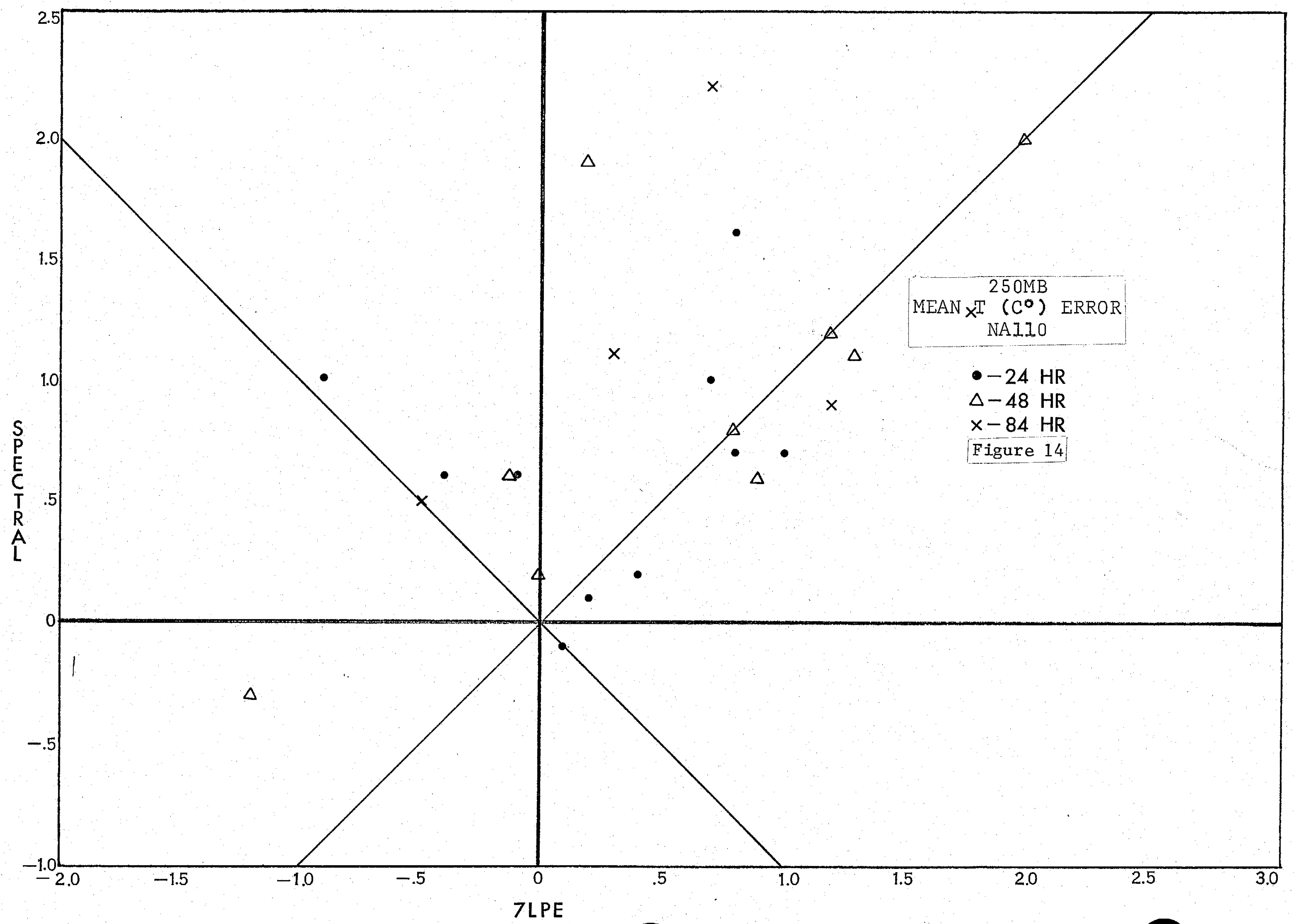
160
140
120
100
80
60
40
20
SPECTRAL
500MB
RMS HT ERROR
NA110
● – 24 HR
△ – 48 HR
× – 84 HR
SPECTRAL
BETTER
7LPE

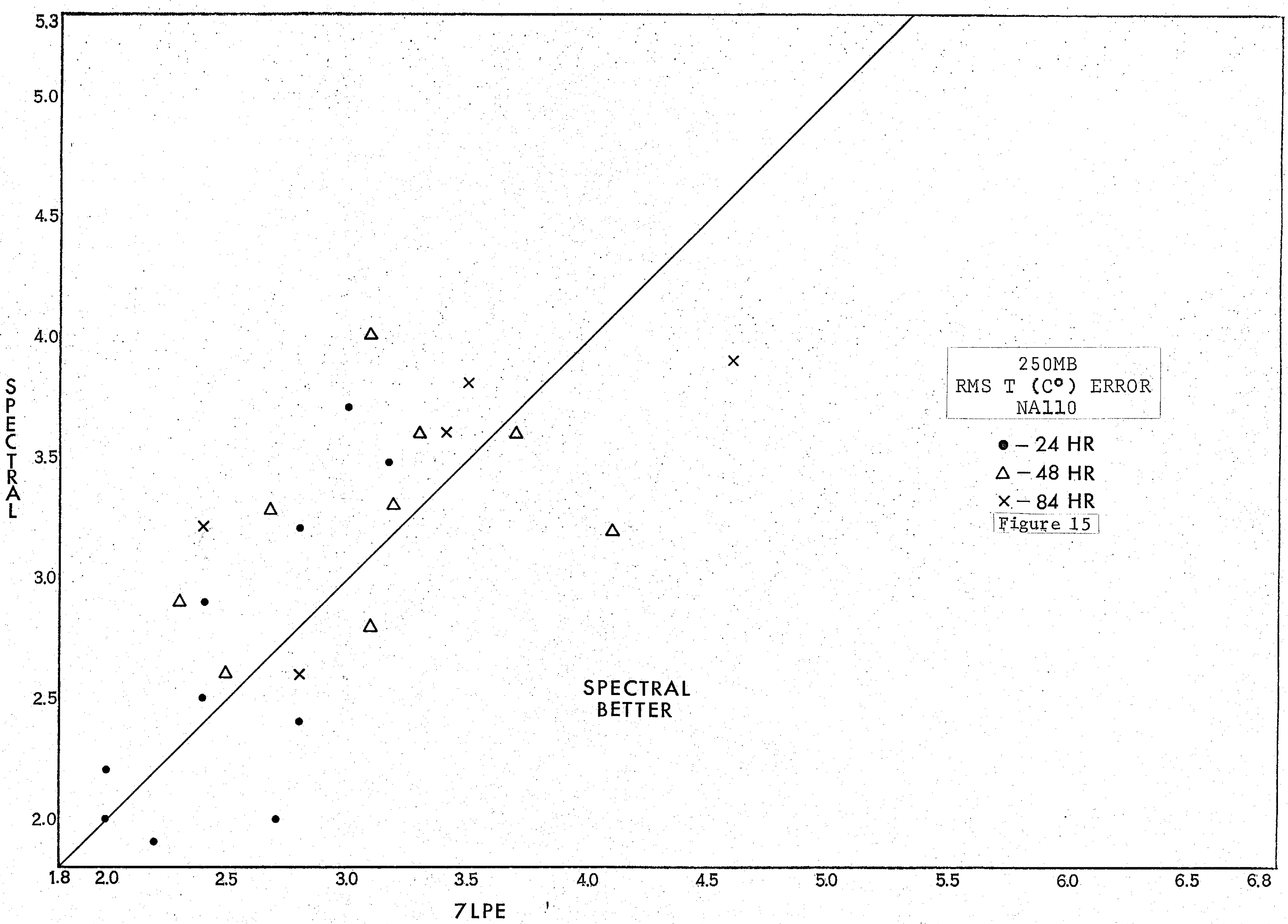
Figure 12



160
140
120
100
80
60
40
20
40
60
80
100
120
140
160
SPECTRAL
7LPE
250MB
RMS HT ERROR
NA110
● – 24 HR
△ – 48 HR
× – 84 HR
SPECTRAL
BETTER

Figure 13


250MB
MEAN T (C°) ERROR
NA110
• – 24 HR
△ – 48 HR
× – 84 HR
SPECTRAL
7LPE
2.5
2.0
1.5
1.0
.5
0
–.5
–1.0
–2.0
–1.5
–1.0
–.5
0
.5
1.0
1.5
2.0
2.5
3.0
Figure 14


250MB
RMS T (C°) ERROR
NA110
● – 24 HR
△ – 48 HR
× – 84 HR
SPECTRAL BETTER
SPECTRAL
7LPE
5.3
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.8
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
6.8
Figure 15

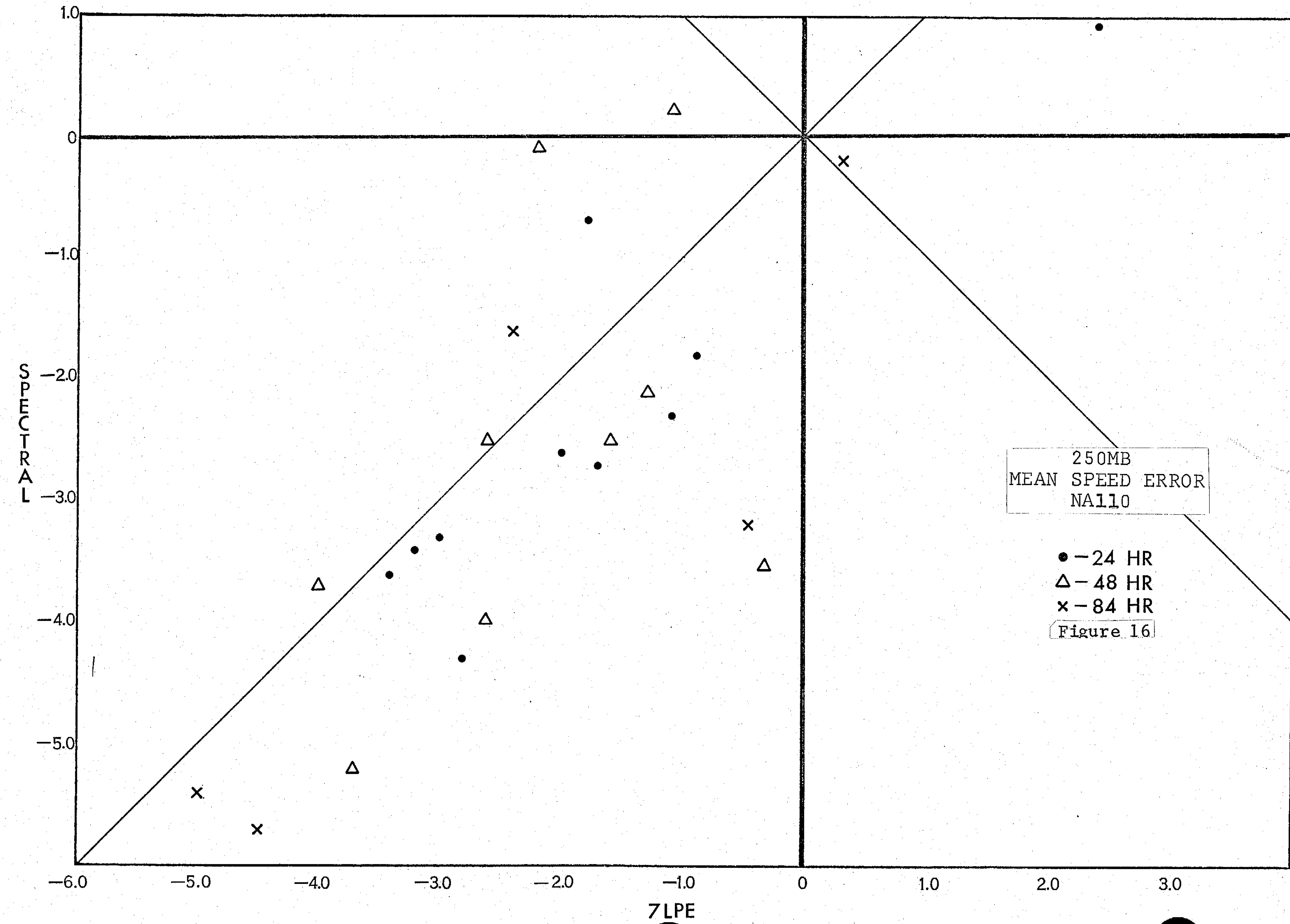

1.0
0
−1.0
−2.0
−3.0
−4.0
−5.0
SPECTRAL
−6.0
−5.0
−4.0
−3.0
−2.0
−1.0
0
1.0
2.0
3.0
7LPE
250MB
MEAN SPEED ERROR
NA110
• − 24 HR
△ − 48 HR
× − 84 HR

Figure 16

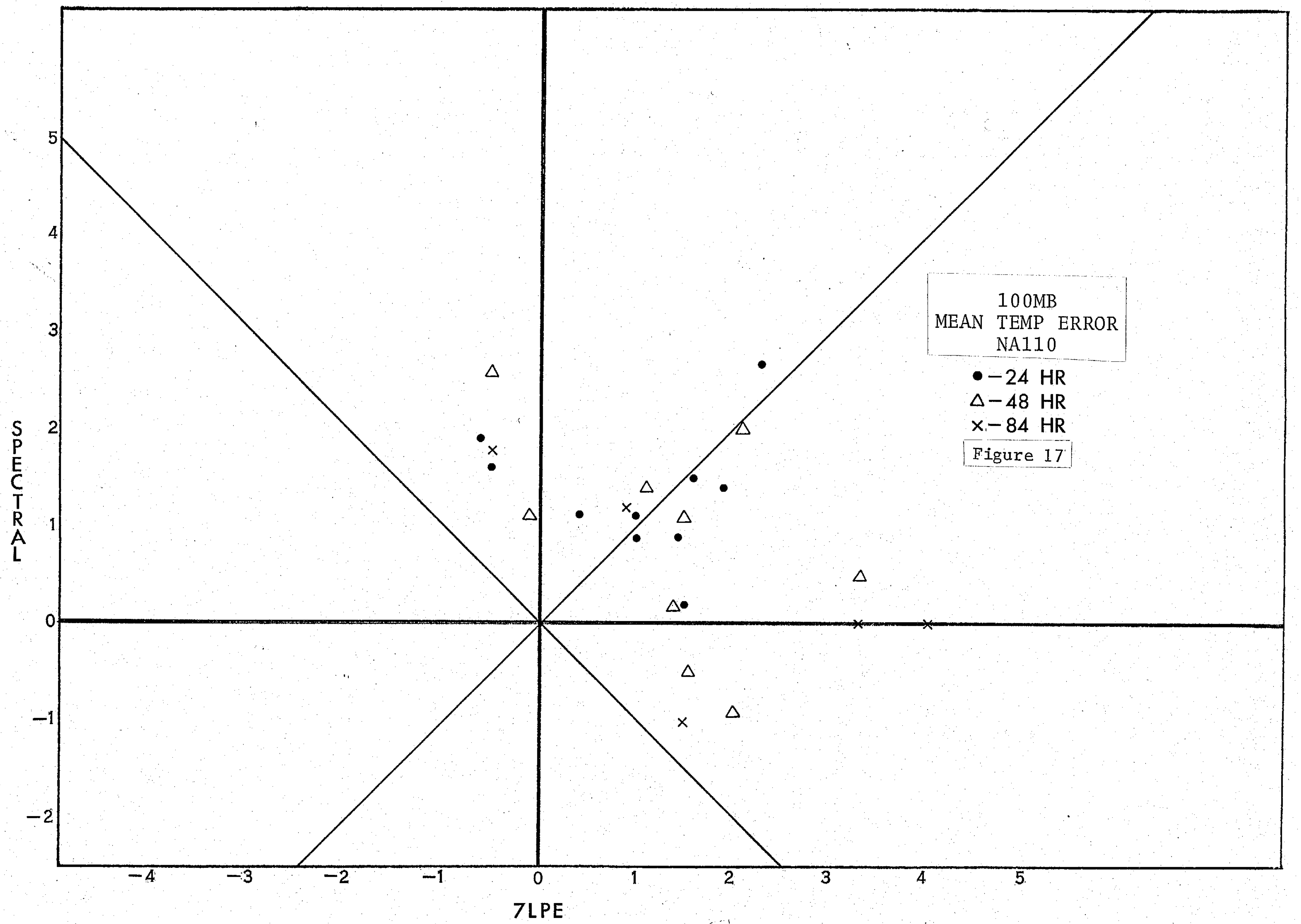
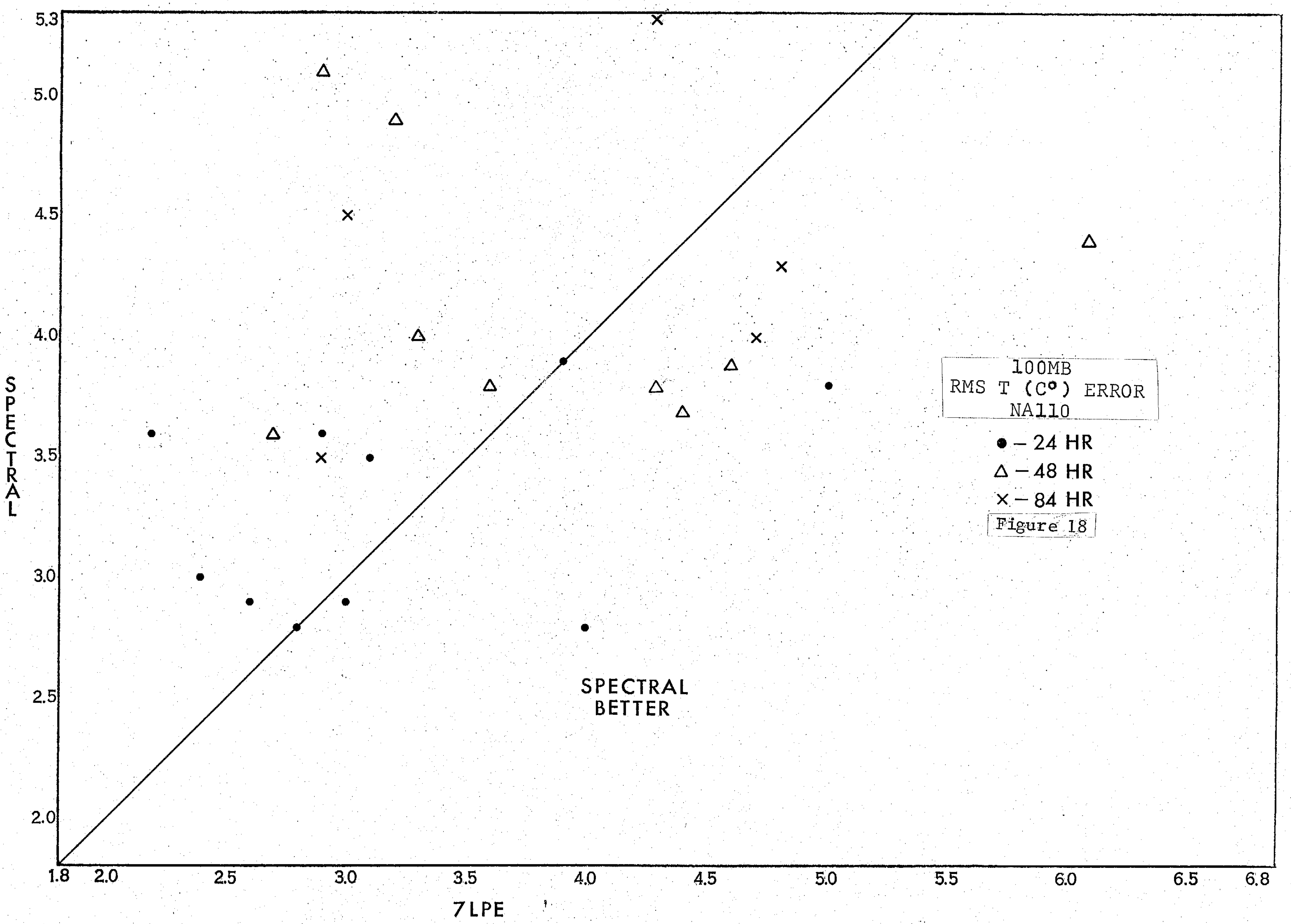
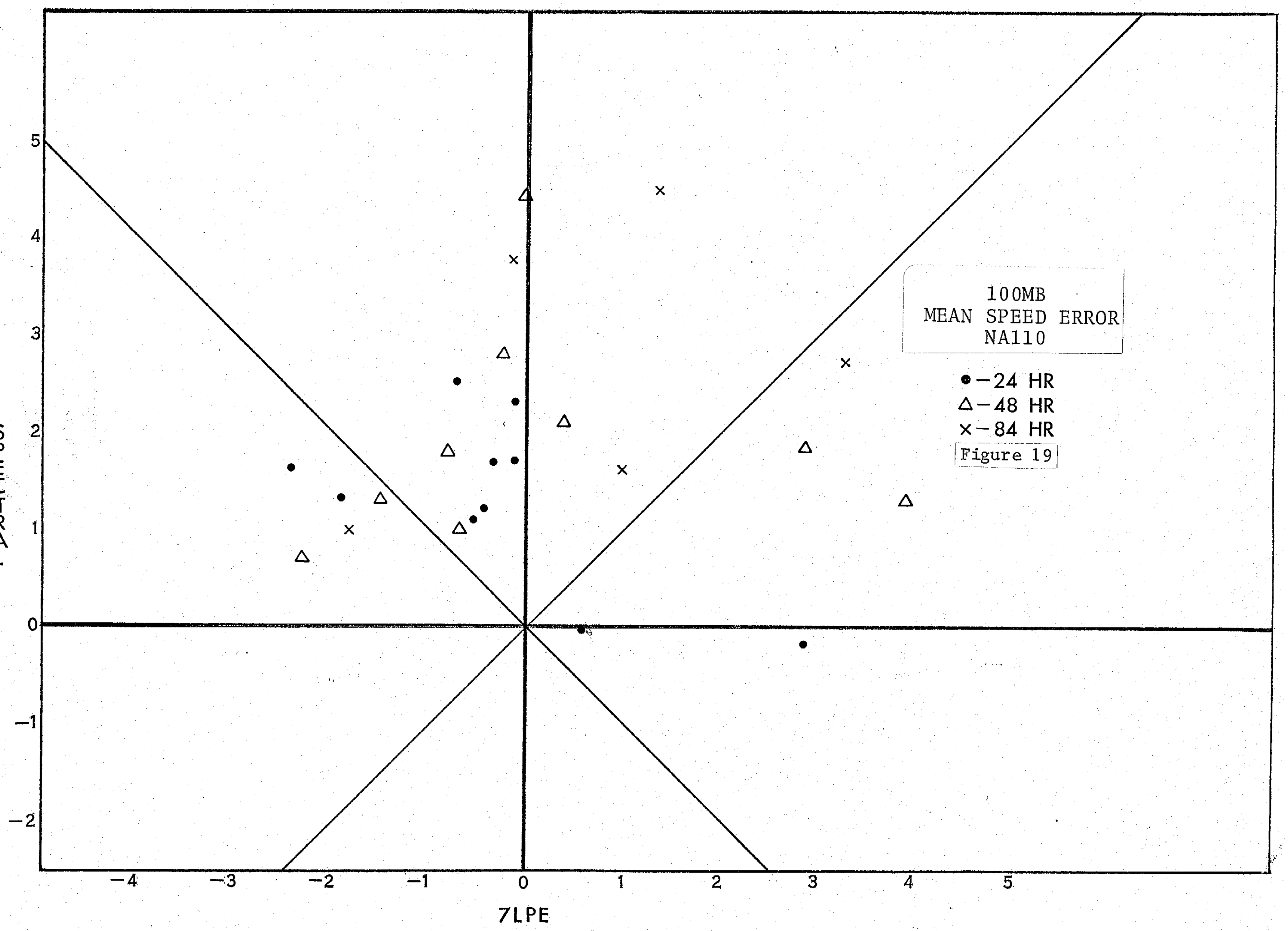

100MB
MEAN TEMP ERROR
NA110
● – 24 HR
△ – 48 HR
× – 84 HR
SPECTRAL
7LPE
5
4
3
2
1
0
–1
–2
–4
–3
–2
–1
0
1
2
3
4
5

Figure 17


100MB
RMS T (C°) ERROR
NA110
• – 24 HR
△ – 48 HR
× – 84 HR
SPECTRAL
BETTER
SPECTRAL
7LPE
5.3
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.8
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
6.8
Figure 18


100MB
MEAN SPEED ERROR
NA110
• – 24 HR
△ – 48 HR
× – 84 HR
Figure 19
SPECTRAL
7LPE
5
4
3
2
1
0
–1
–2
–4
–3
–2
–1
0
1
2
3
4
5

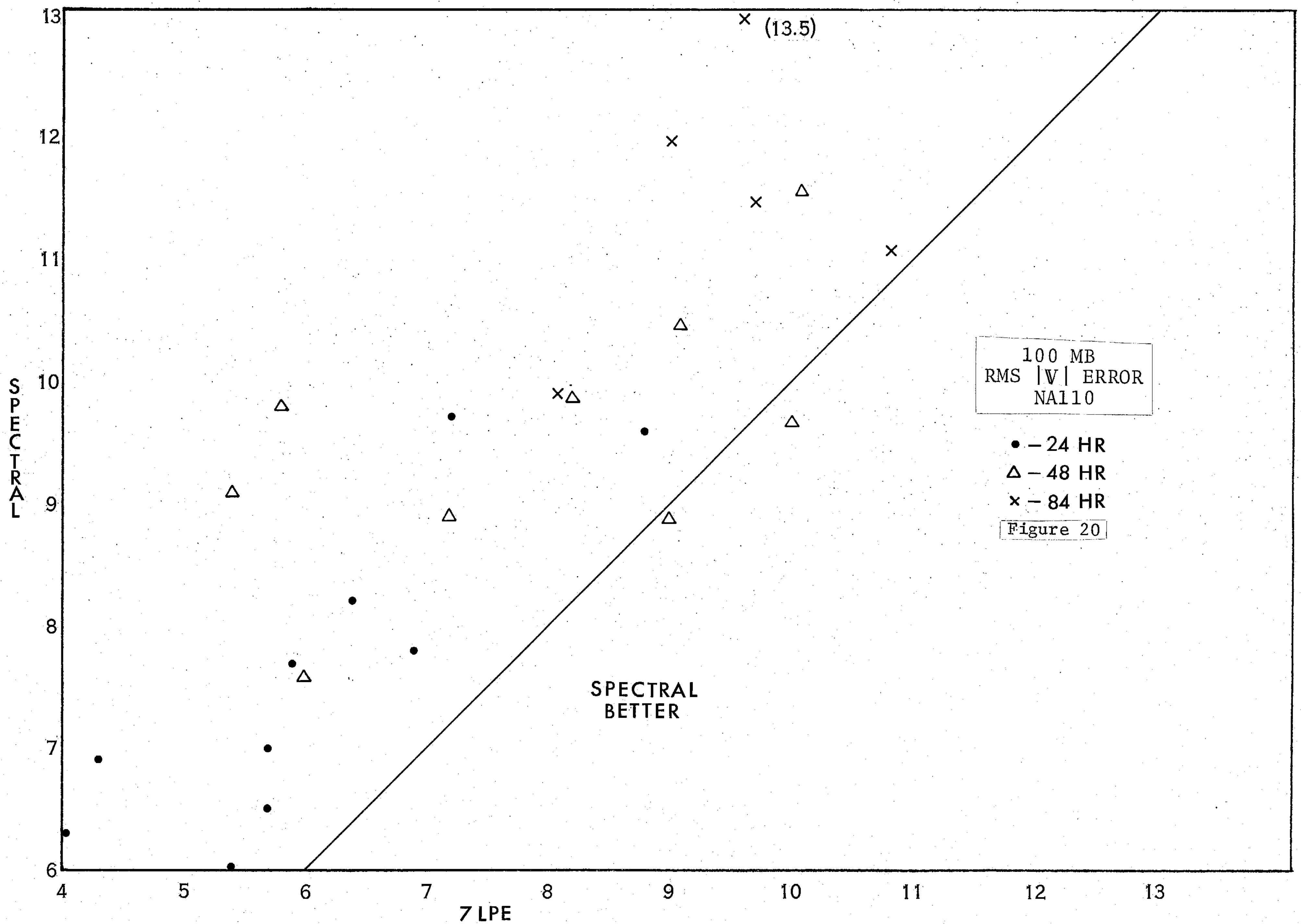
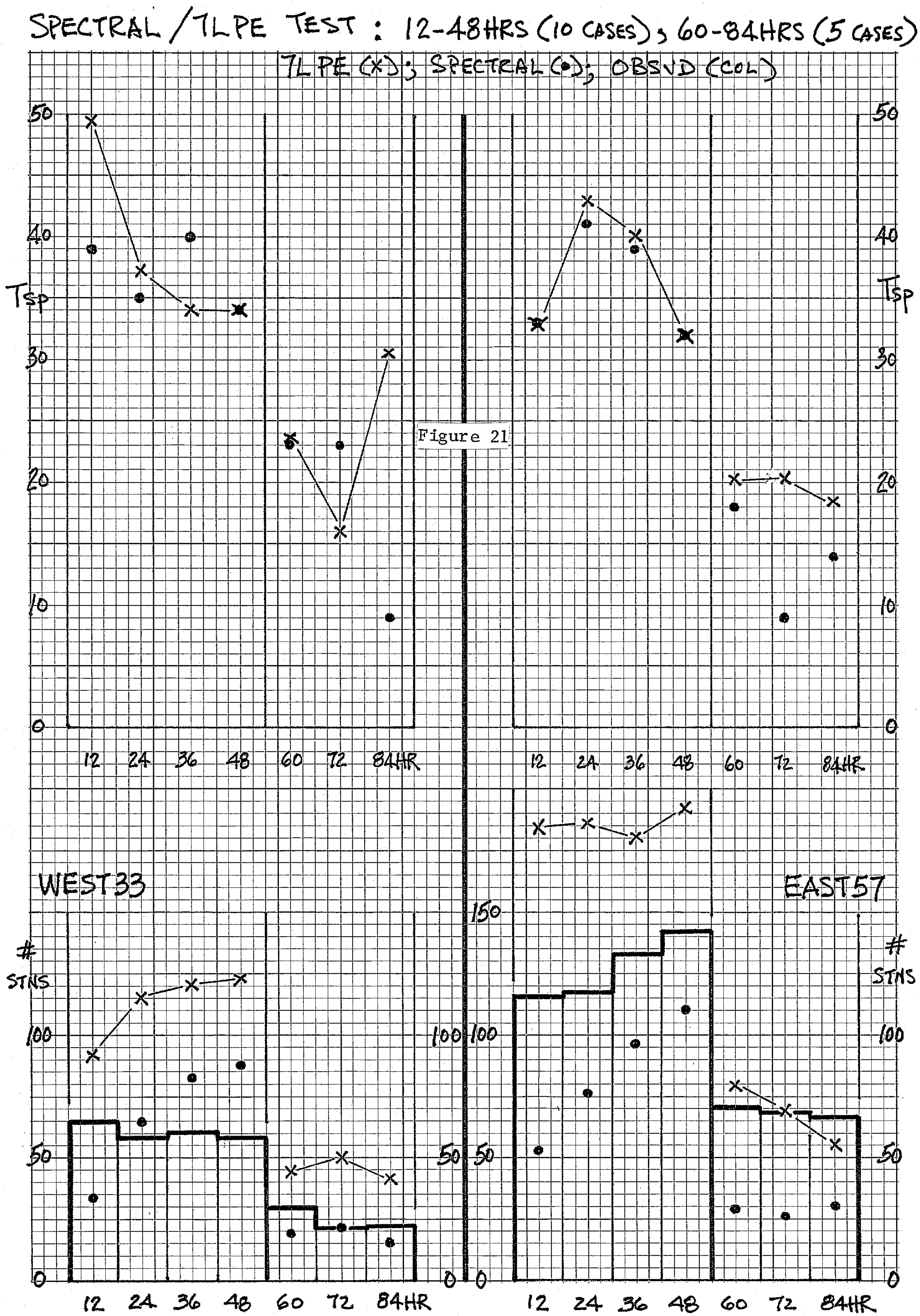

13
12
11
10
9
8
7
6
4
5
6
7
8
9
10
11
12
13
SPECTRAL
7 LPE
(13.5)
SPECTRAL BETTER
100 MB
RMS |V| ERROR
NA110
• – 24 HR
Δ – 48 HR
× – 84 HR

Figure 20


SPECTRAL / 7LPE TEST : 12-48HRS (10 CASES) ; 60-84HRS (5 CASES)
7LPE (X); SPECTRAL (•); OBSVD (COL)
Figure 21
Tsp
50
40
30
20
10
0
12 24 36 48 60 72 84HR
WEST33
EAST57
STNS
150
100
50
FPI-MI-10 X 10 TO 1 INCH
5TH LINE ACCENTED, 10TH HEAVY

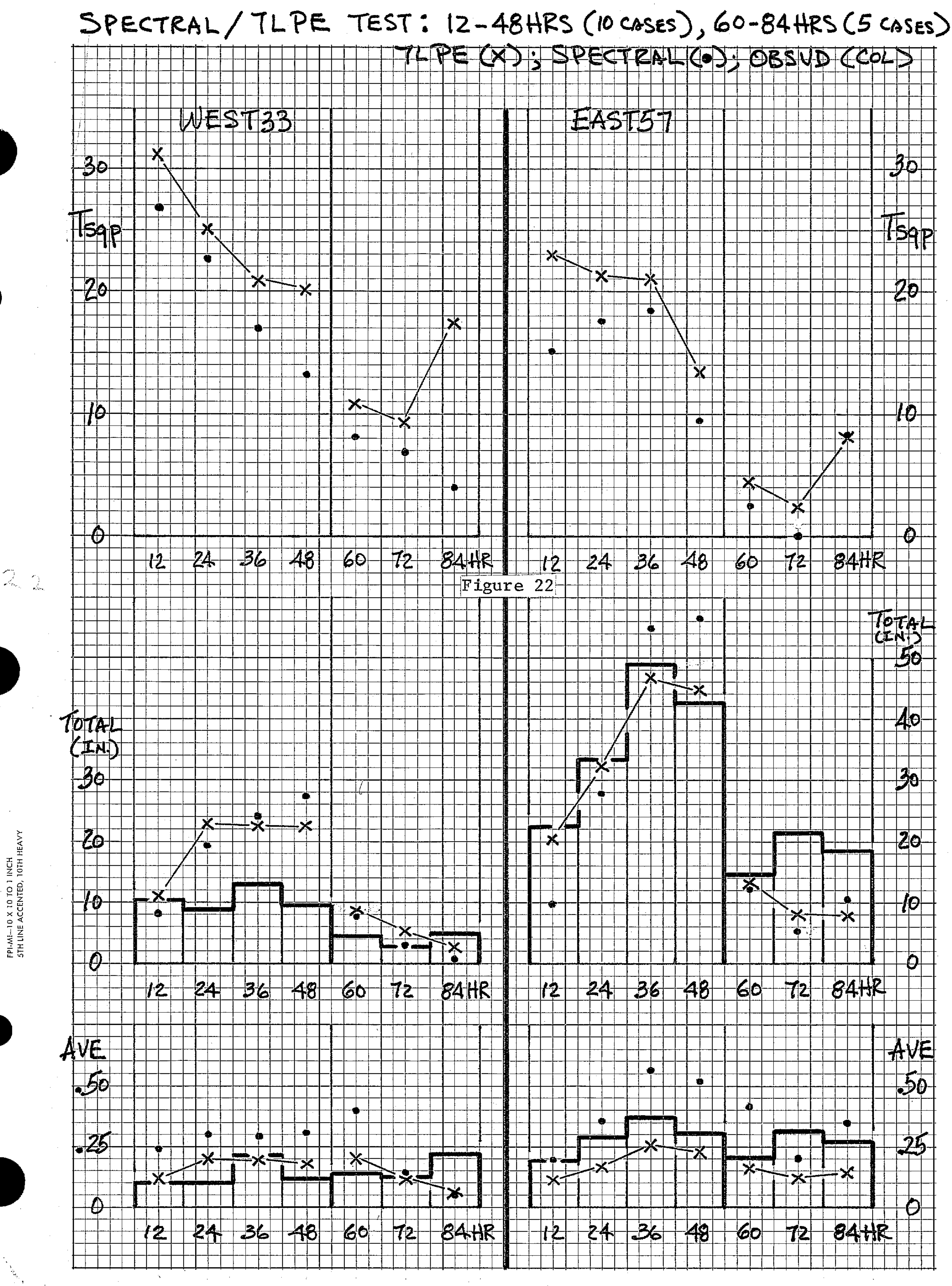
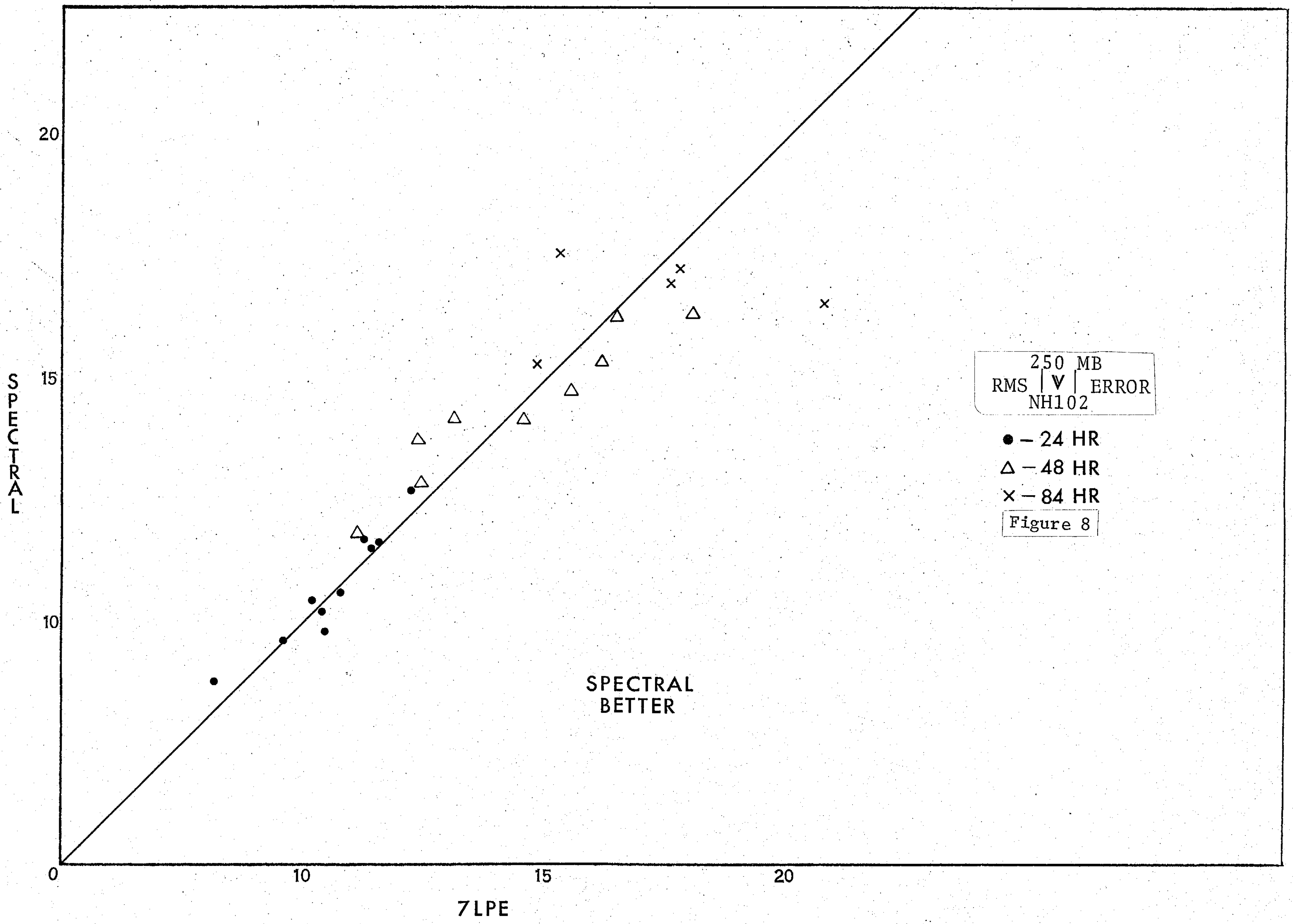
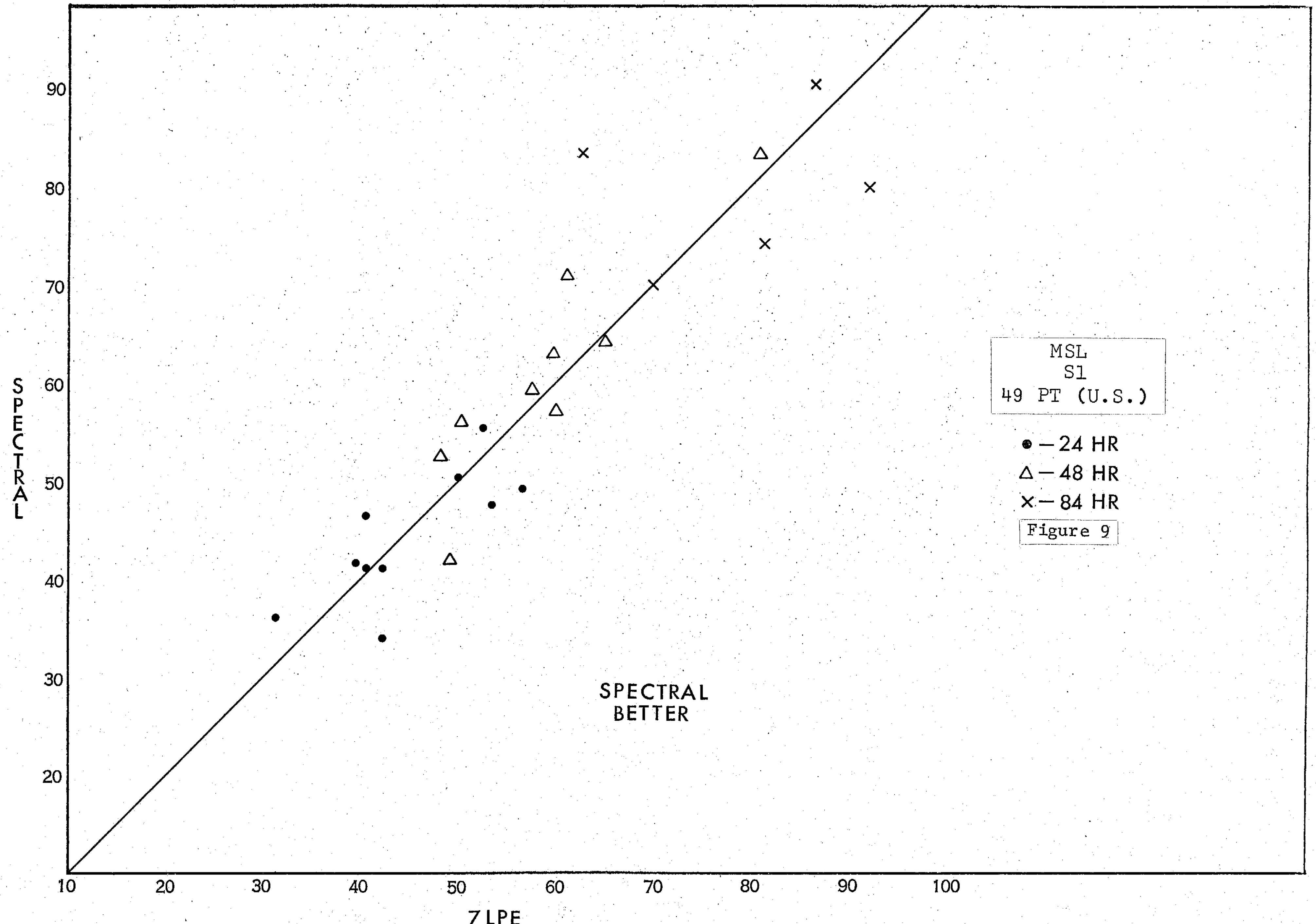

SPECTRAL / TLPE TEST: 12-48HRS (10 CASES), 60-84HRS (5 CASES)
TLPE (X); SPECTRAL (•); OBSVD (COL)
WEST33
EAST57
Tsqp
TOTAL (IN.)
AVE
12 24 36 48 60 72 84HR
FPI-MI-10 X 10 TO 1 INCH
5TH LINE ACCENTED, 10TH HEAVY

Figure 22



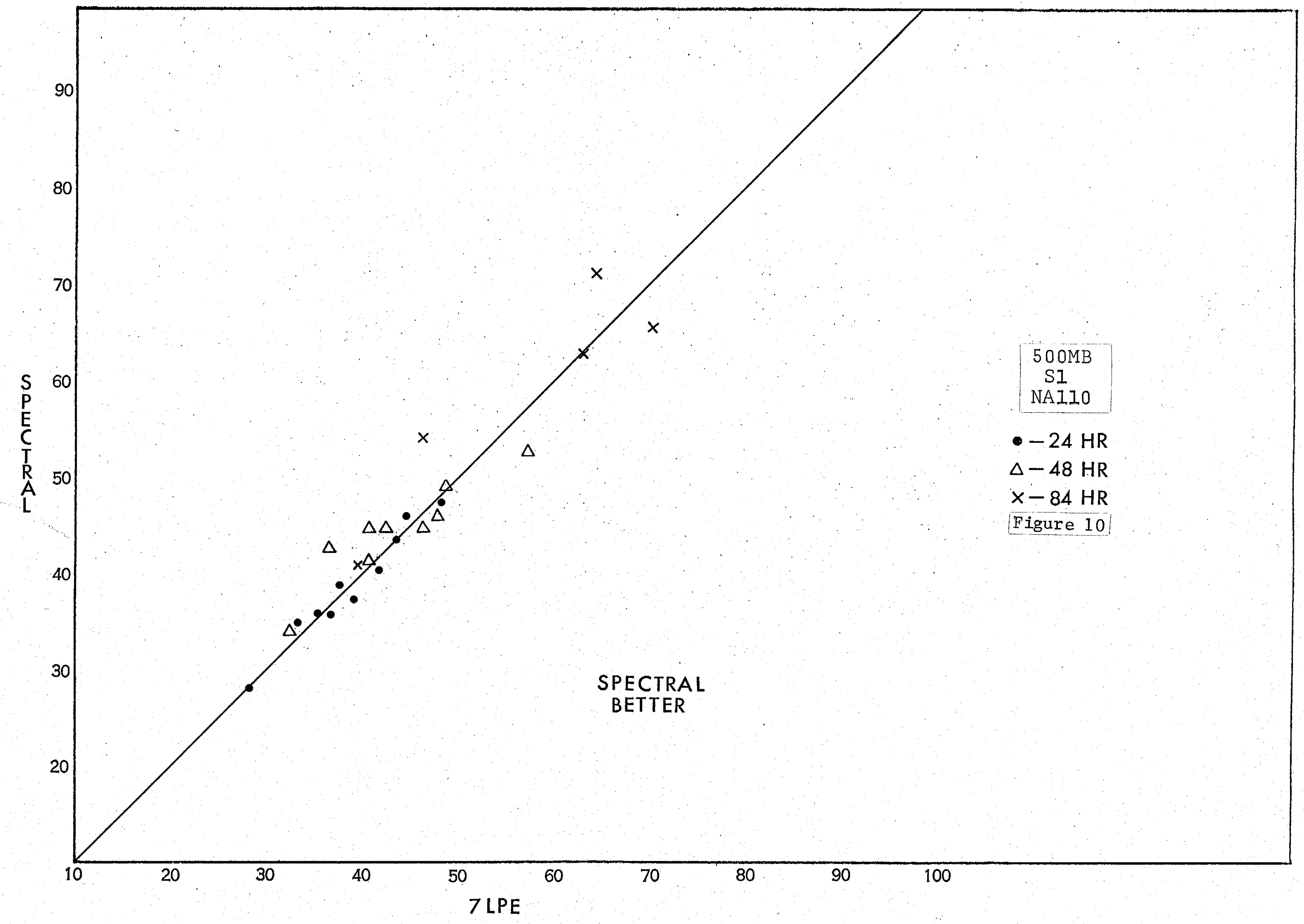
250 MB
RMS |V| ERROR
NH102
● – 24 HR
△ – 48 HR
× – 84 HR
SPECTRAL
BETTER
SPECTRAL
7LPE
0
10
15
20

Figure 8



SPECTRAL
7 LPE
MSL
S1
49 PT (U.S.)
— 24 HR
— 48 HR
— 84 HR
SPECTRAL BETTER
10
20
30
40
50
60
70
80
90
100

Figure 9


500MB
S1
NA110
• — 24 HR
△ — 48 HR
× — 84 HR
SPECTRAL
BETTER
SPECTRAL
7 LPE
90
80
70
60
50
40
30
20
10
20
30
40
50
60
70
80
90
100
Figure 10

MODEL STRUCTURE

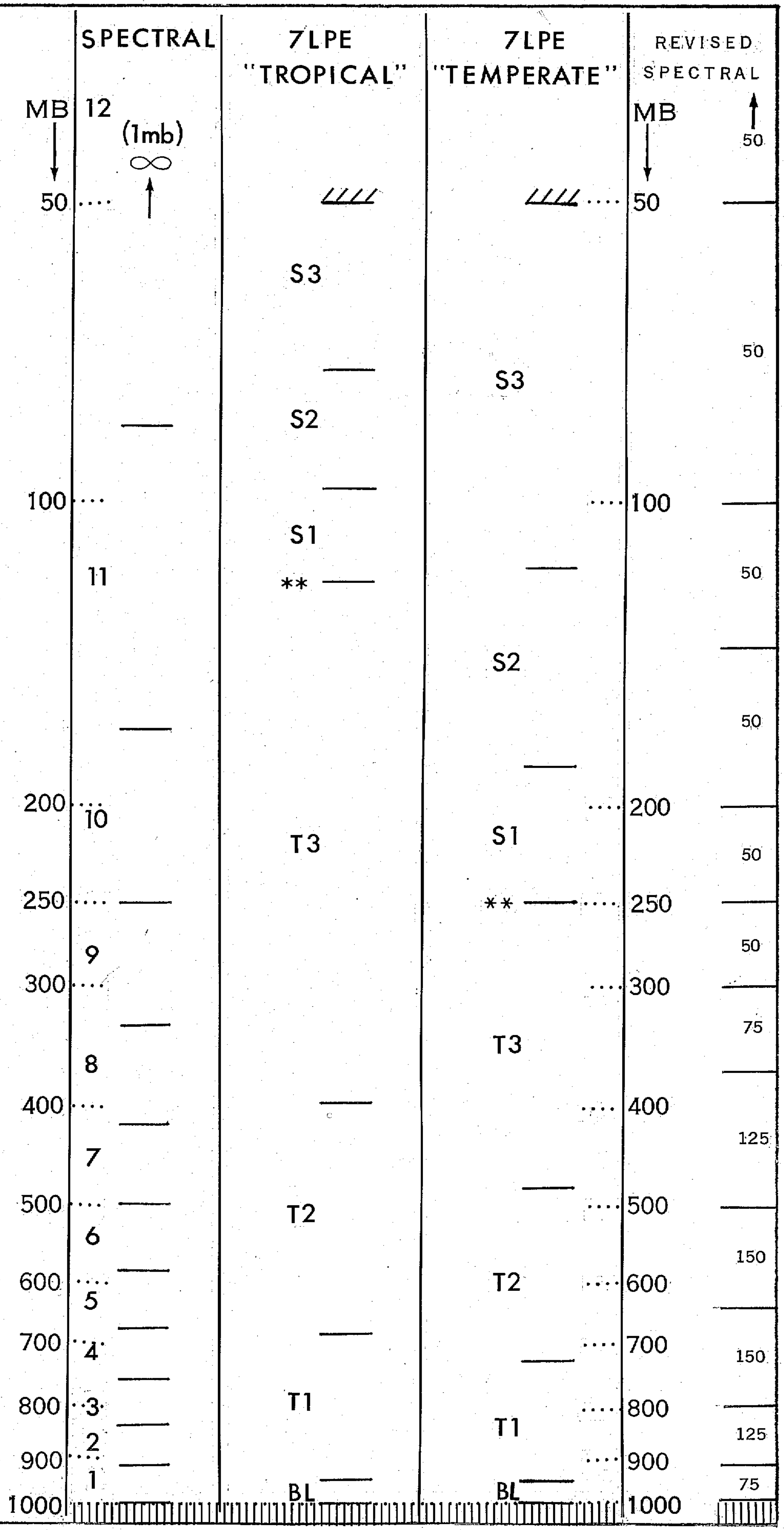
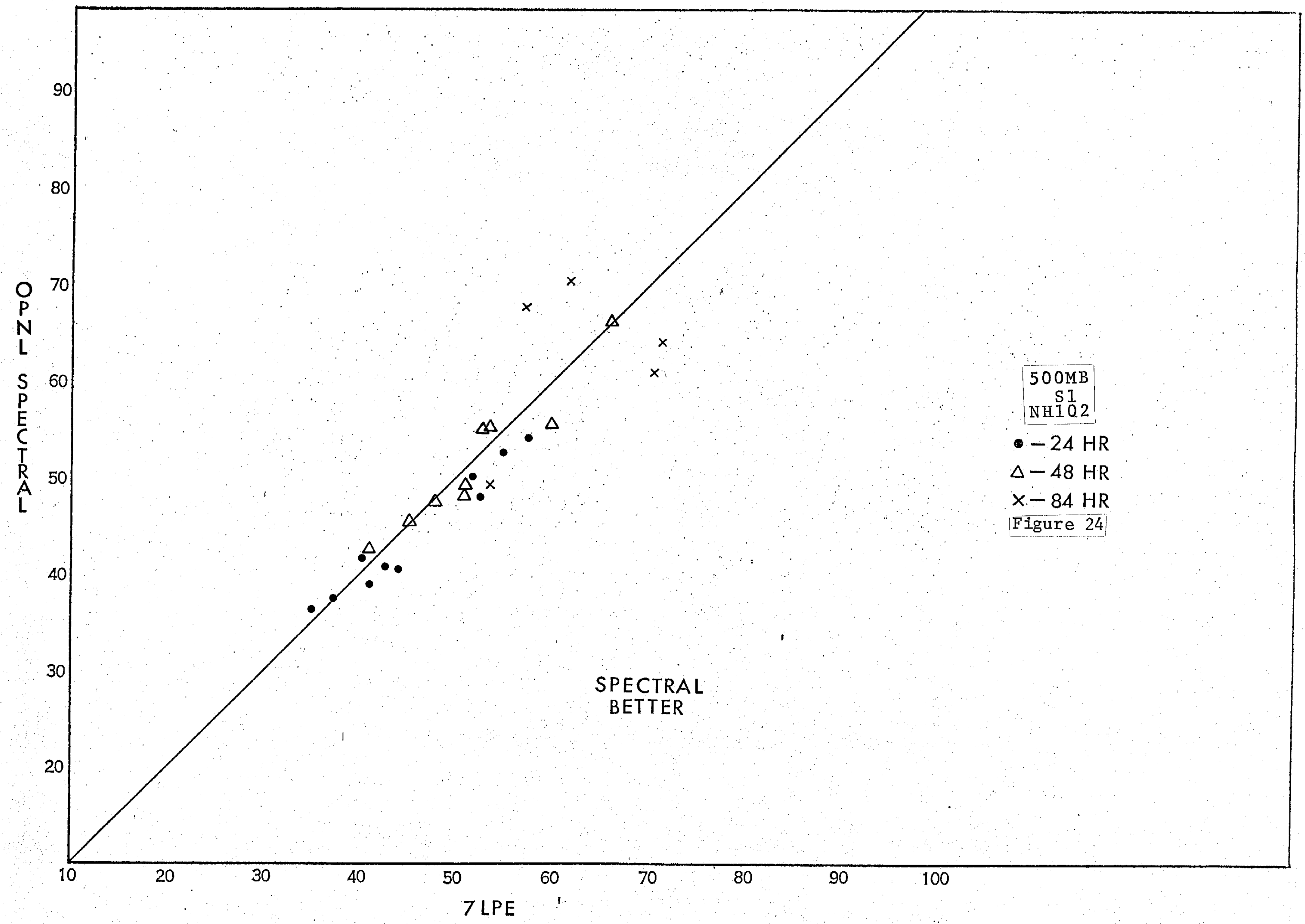
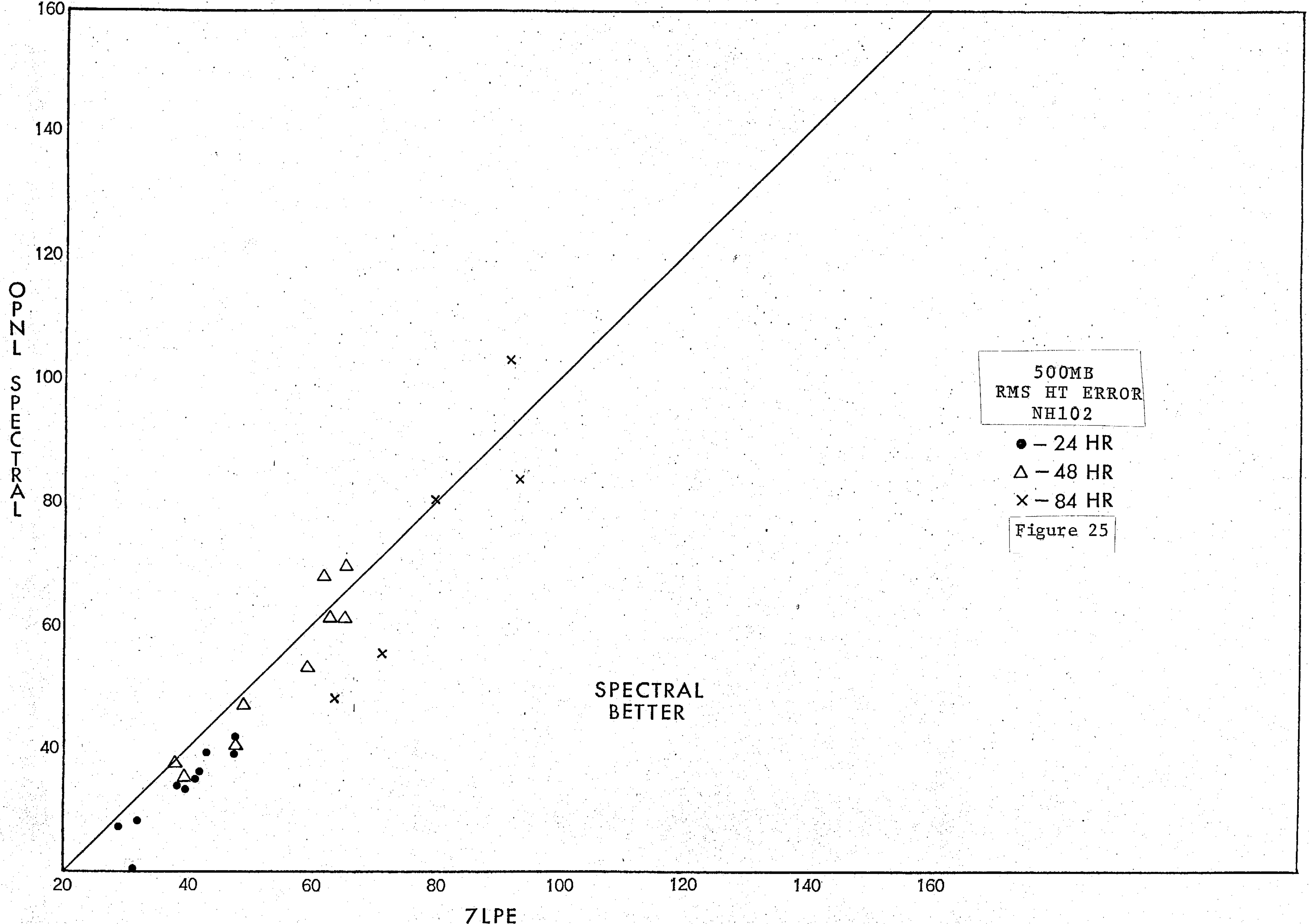


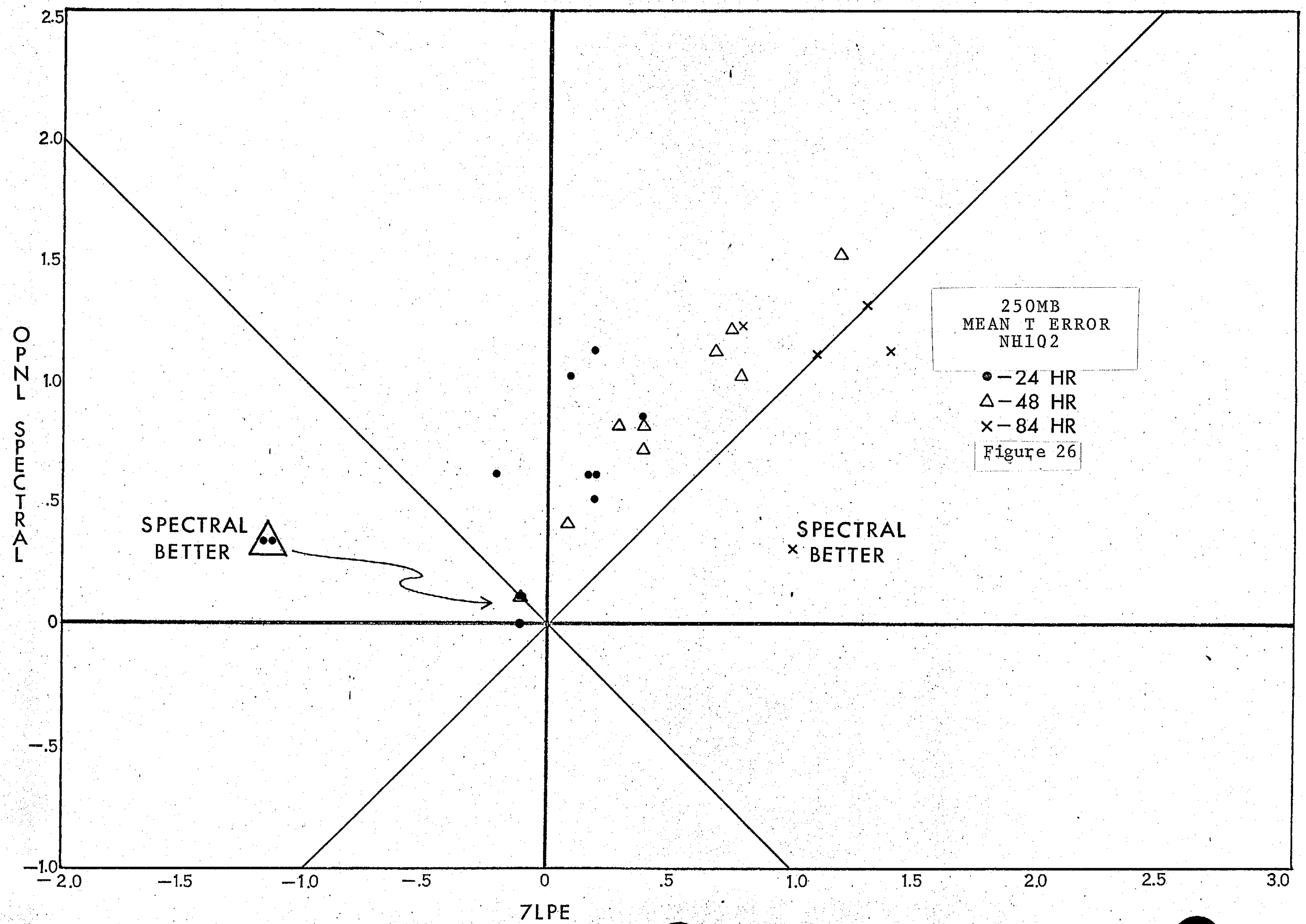
Figure 23

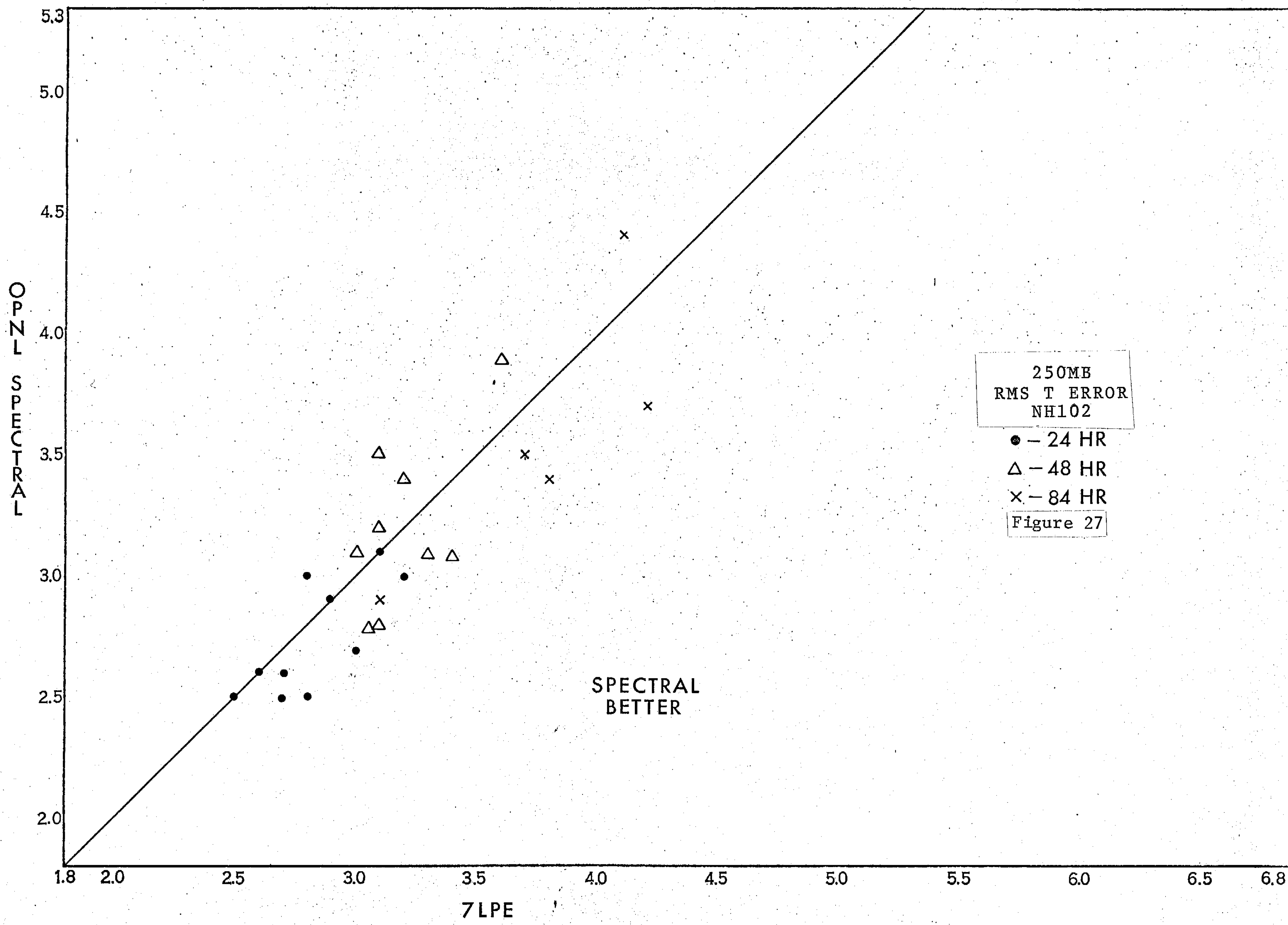


OPNL SPECTRAL
7 LPE
SPECTRAL BETTER
500MB
S1
NH102
— 24 HR
— 48 HR
— 84 HR
10
20
30
40
50
60
70
80
90
100

Figure 24

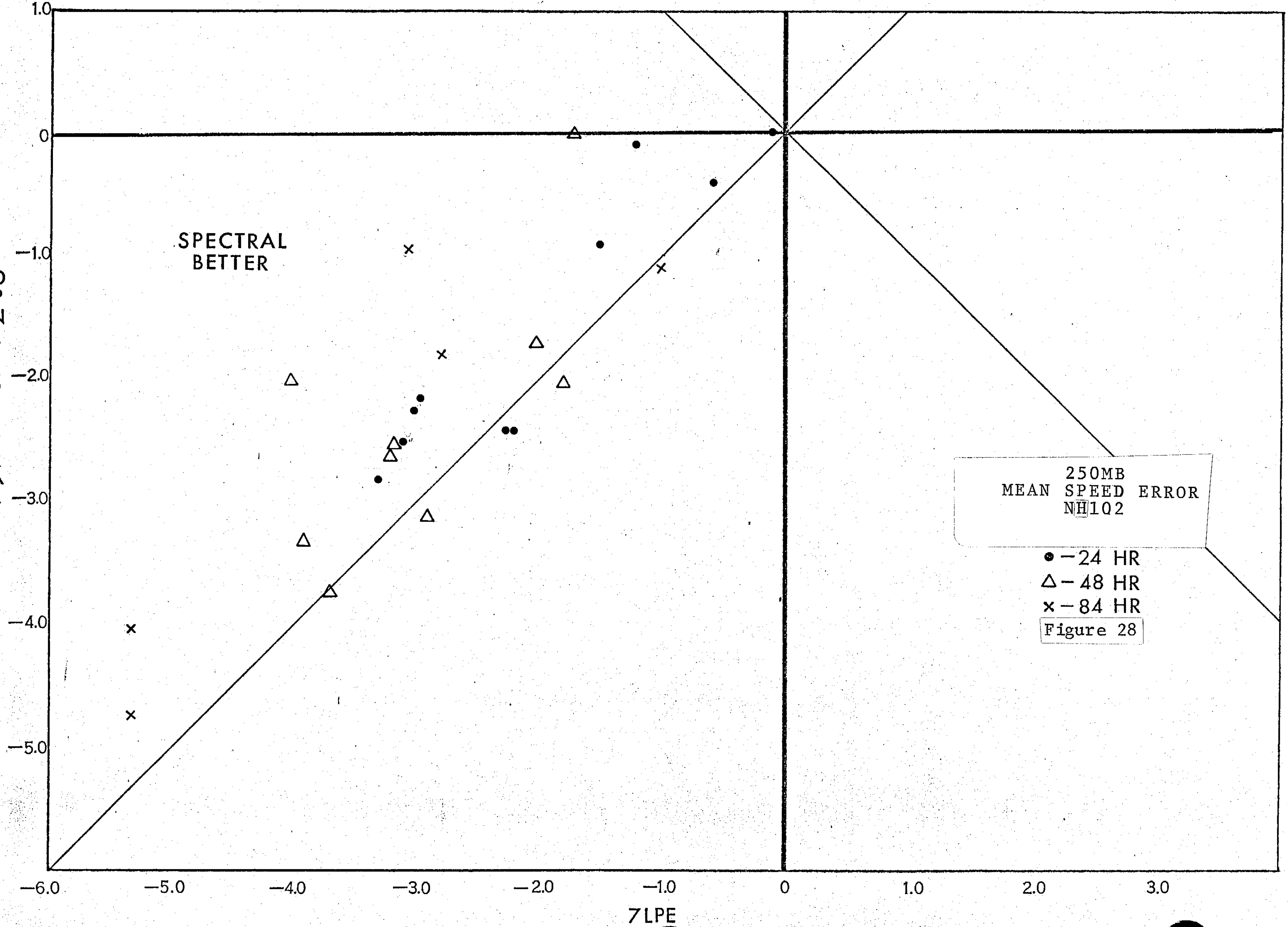

500MB
RMS HT ERROR
NH102
● – 24 HR
△ – 48 HR
× – 84 HR
SPECTRAL BETTER
OPNL SPECTRAL
7LPE
160
140
120
100
80
60
40
20
40
60
80
100
120
140
160
Figure 25


250MB
MEAN T ERROR
NH102
–24 HR
–48 HR
–84 HR
Figure 26
SPECTRAL BETTER
SPECTRAL BETTER
OPNL SPECTRAL
7LPE
2.5
2.0
1.5
1.0
.5
0
−.5
−1.0
−2.0
−1.5
−1.0
−.5
0
.5
1.0
1.5
2.0
2.5
3.0



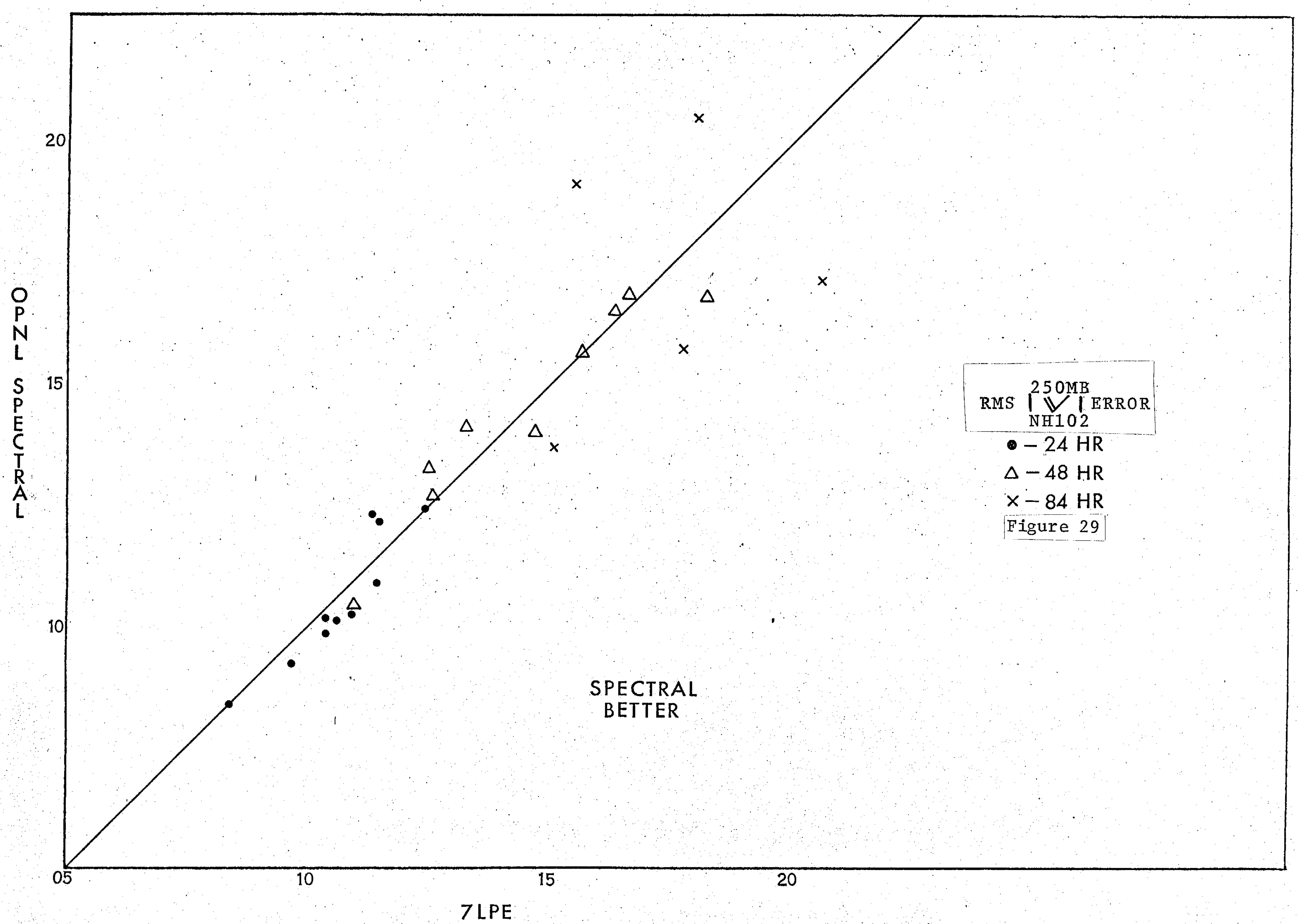
OPNL SPECTRAL
7LPE
SPECTRAL BETTER
250MB
RMS T ERROR
NH102
– 24 HR
– 48 HR
– 84 HR
5.3
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.8
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
6.8

Figure 27



1.0
0
-1.0
-2.0
-3.0
-4.0
-5.0
OPNL SPECTRAL
SPECTRAL BETTER
250MB
MEAN SPEED ERROR
NH102
• - 24 HR
△ - 48 HR
× - 84 HR
-6.0
-5.0
-4.0
-3.0
-2.0
-1.0
0
1.0
2.0
3.0
7LPE

Figure 28



OPNL SPECTRAL
20
15
10
05
10
15
20
7LPE
SPECTRAL BETTER
250MB
RMS | V | ERROR
NH102
– 24 HR
– 48 HR
– 84 HR

Figure 29

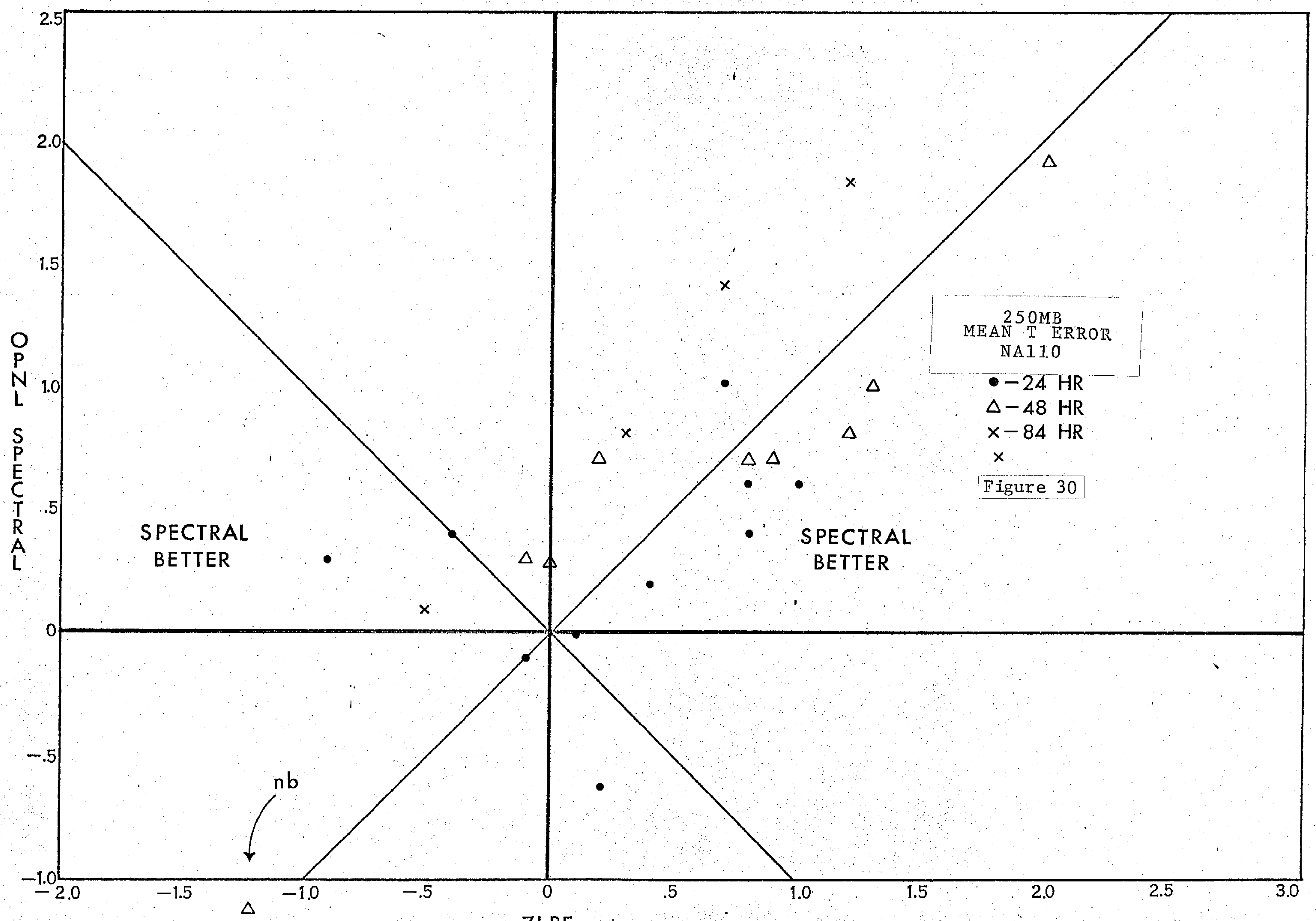
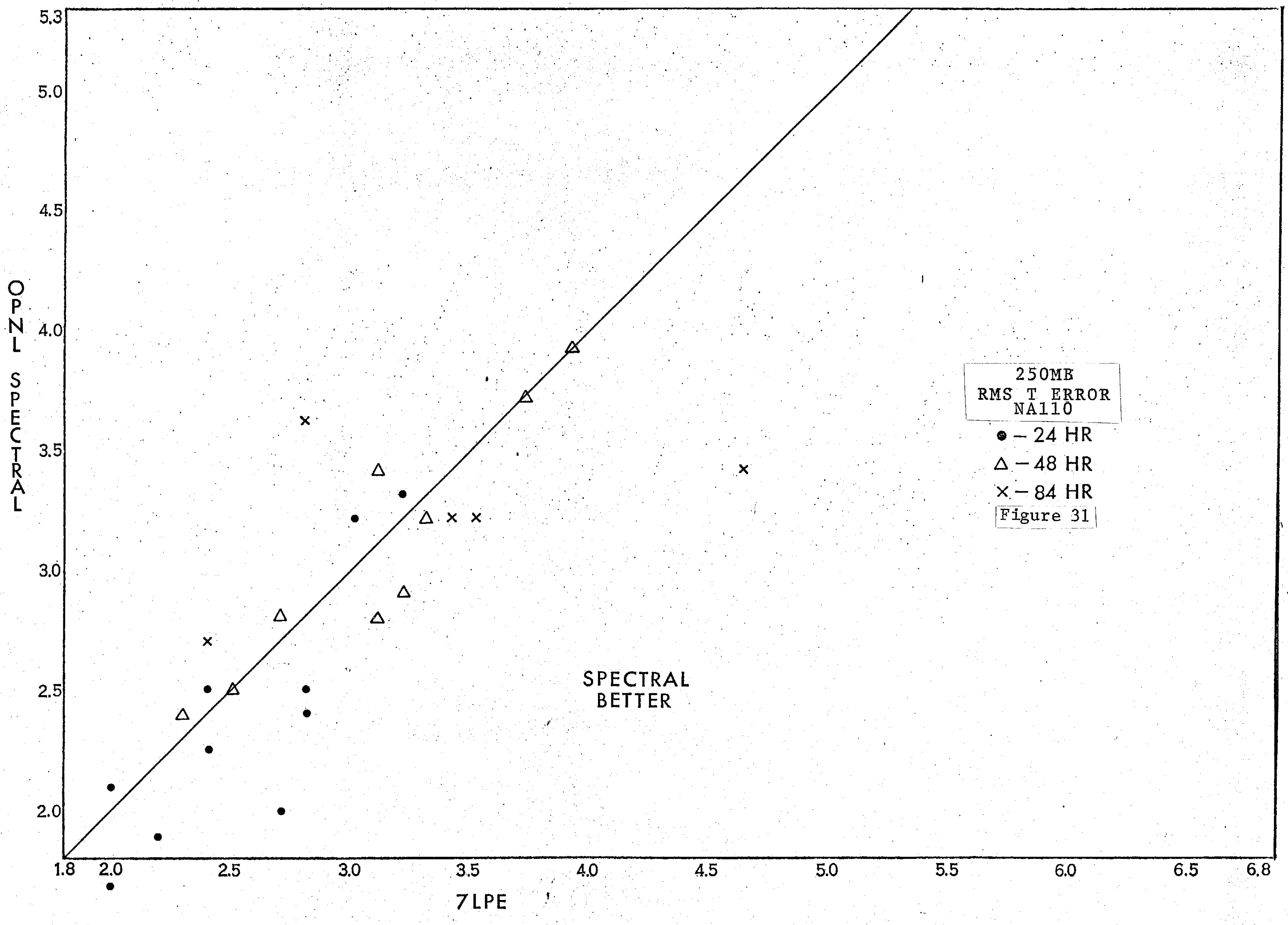
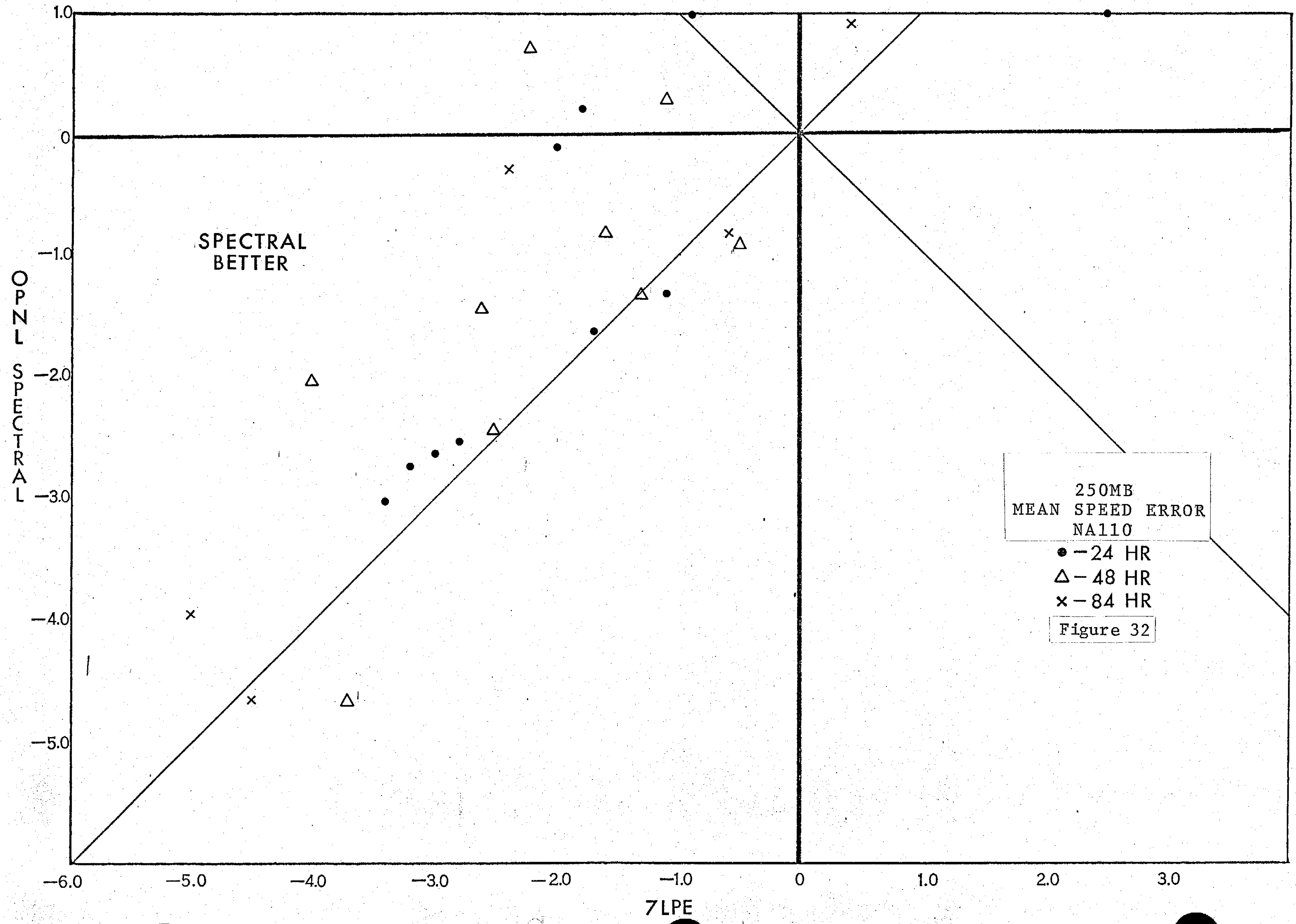

250MB
MEAN T ERROR
NA110
● − 24 HR
△ − 48 HR
× − 84 HR
SPECTRAL BETTER
SPECTRAL BETTER
nb
OPNL SPECTRAL
7LPE
2.5
2.0
1.5
1.0
.5
0
−.5
−1.0
− 2.0
−1.5
−1.0
−.5
0
.5
1.0
1.5
2.0
2.5
3.0

Figure 30



250MB
RMS T ERROR
NA110
• – 24 HR
△ – 48 HR
× – 84 HR
SPECTRAL BETTER
OPNL SPECTRAL
7LPE
5.3
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.8
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
6.8

Figure 31



SPECTRAL
BETTER
OPNL SPECTRAL
7LPE
1.0
0
−1.0
−2.0
−3.0
−4.0
−5.0
−6.0
−5.0
−4.0
−3.0
−2.0
−1.0
0
1.0
2.0
3.0
250MB
MEAN SPEED ERROR
NA110
• − 24 HR
△ − 48 HR
× − 84 HR

Figure 32

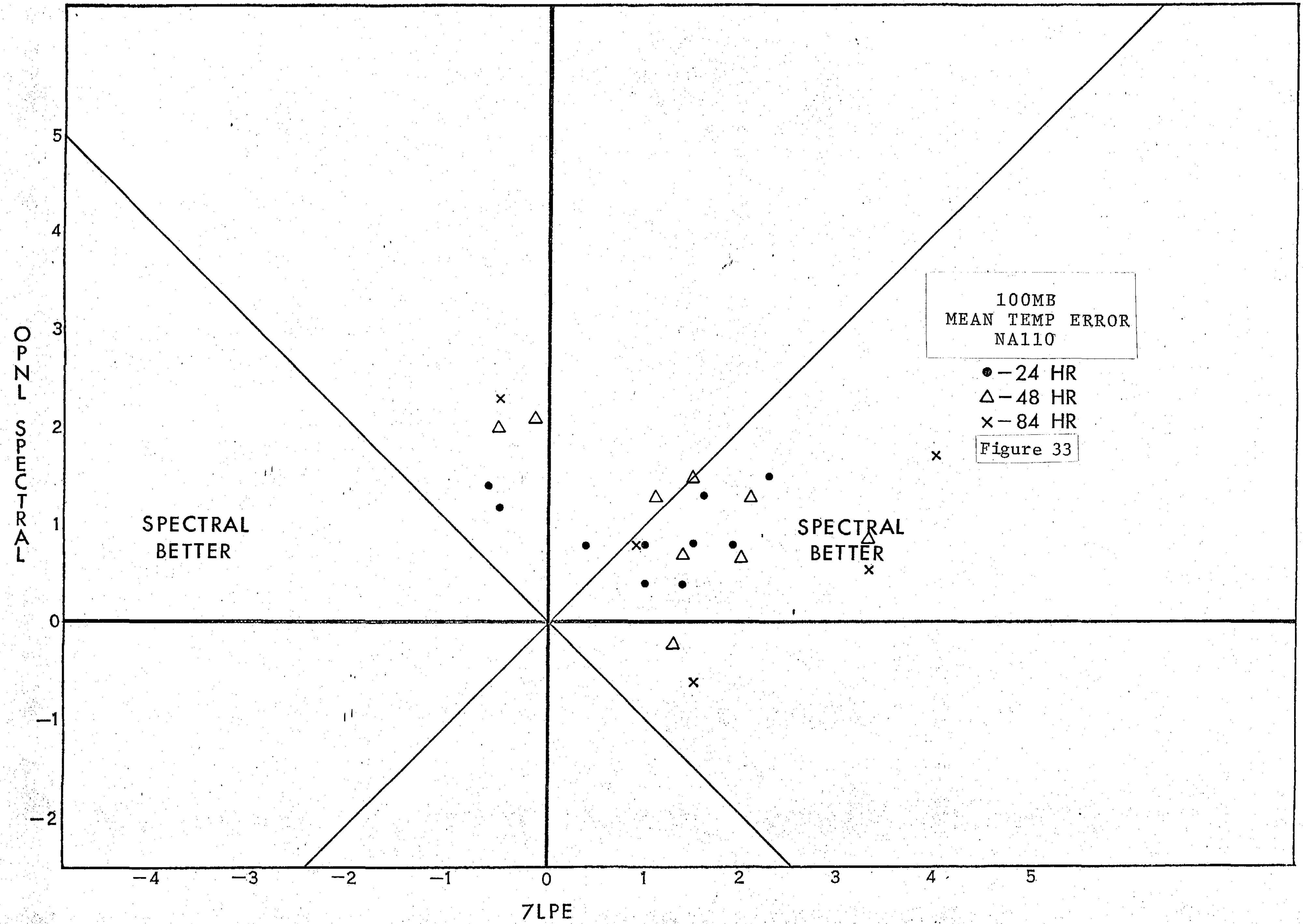
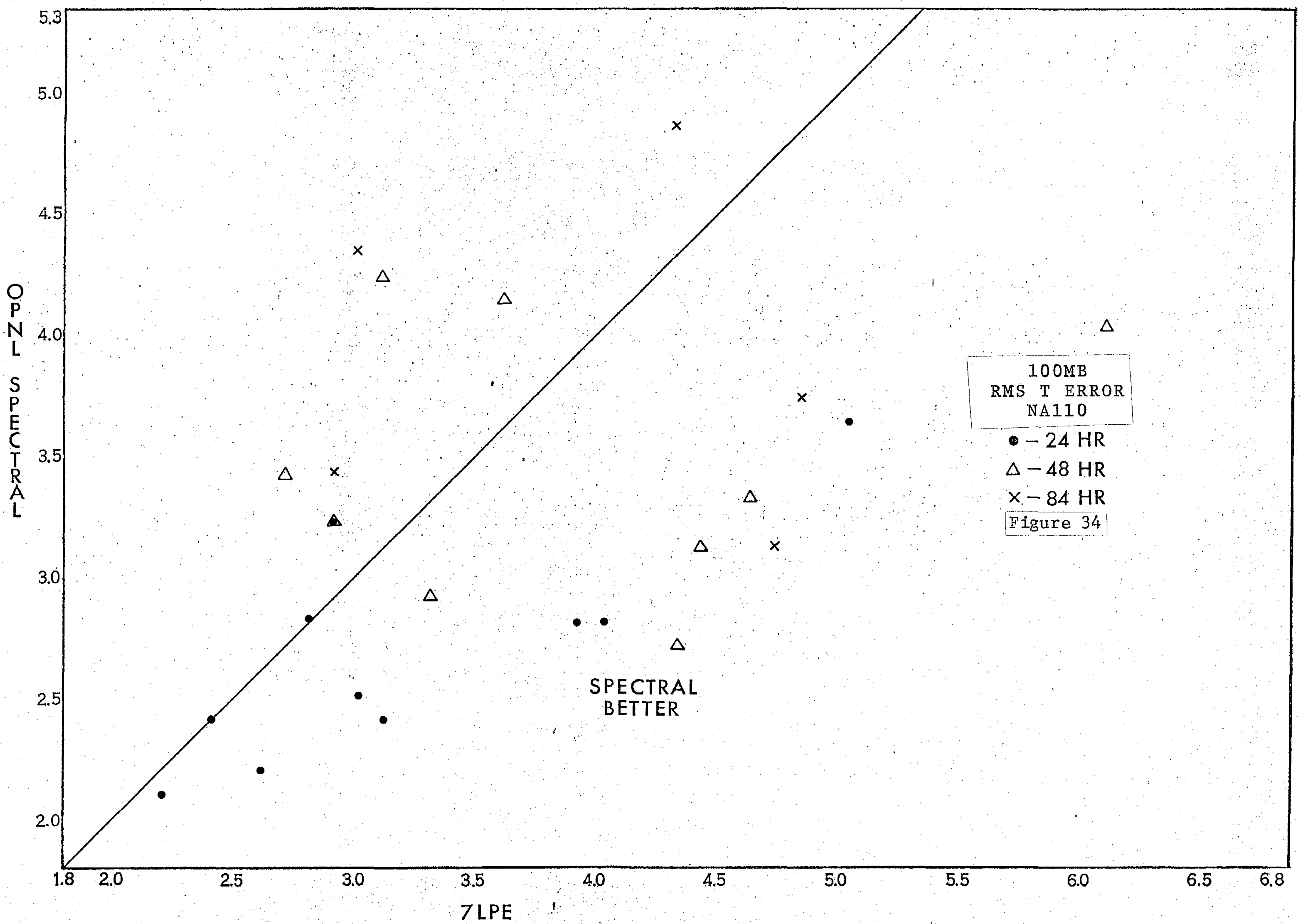
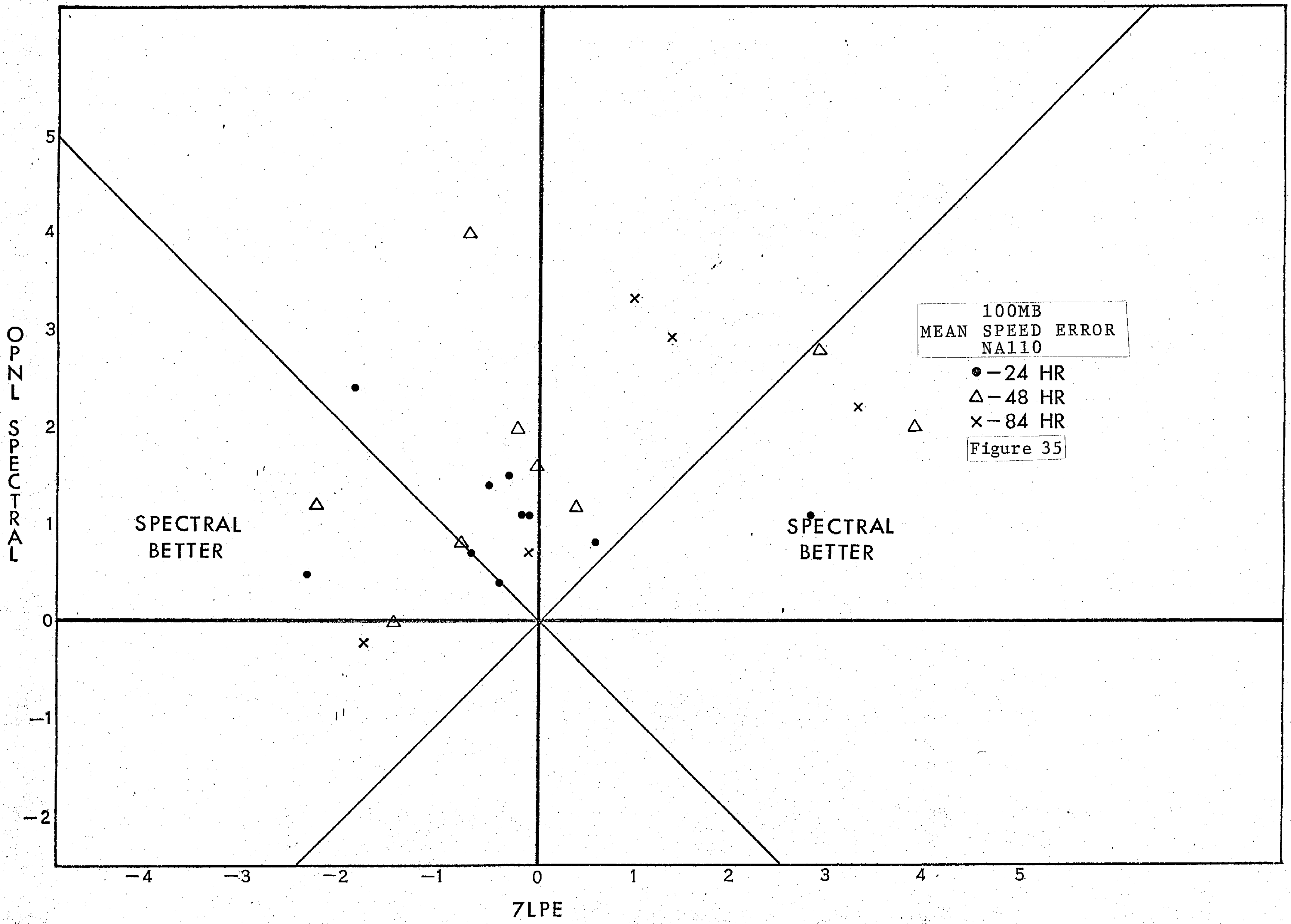

100MB
MEAN TEMP ERROR
NA110
• – 24 HR
△ – 48 HR
× – 84 HR
SPECTRAL BETTER
SPECTRAL BETTER
OPNL SPECTRAL
7LPE
5
4
3
2
1
0
–1
–2
–4
–3
–2
–1
0
1
2
3
4
5

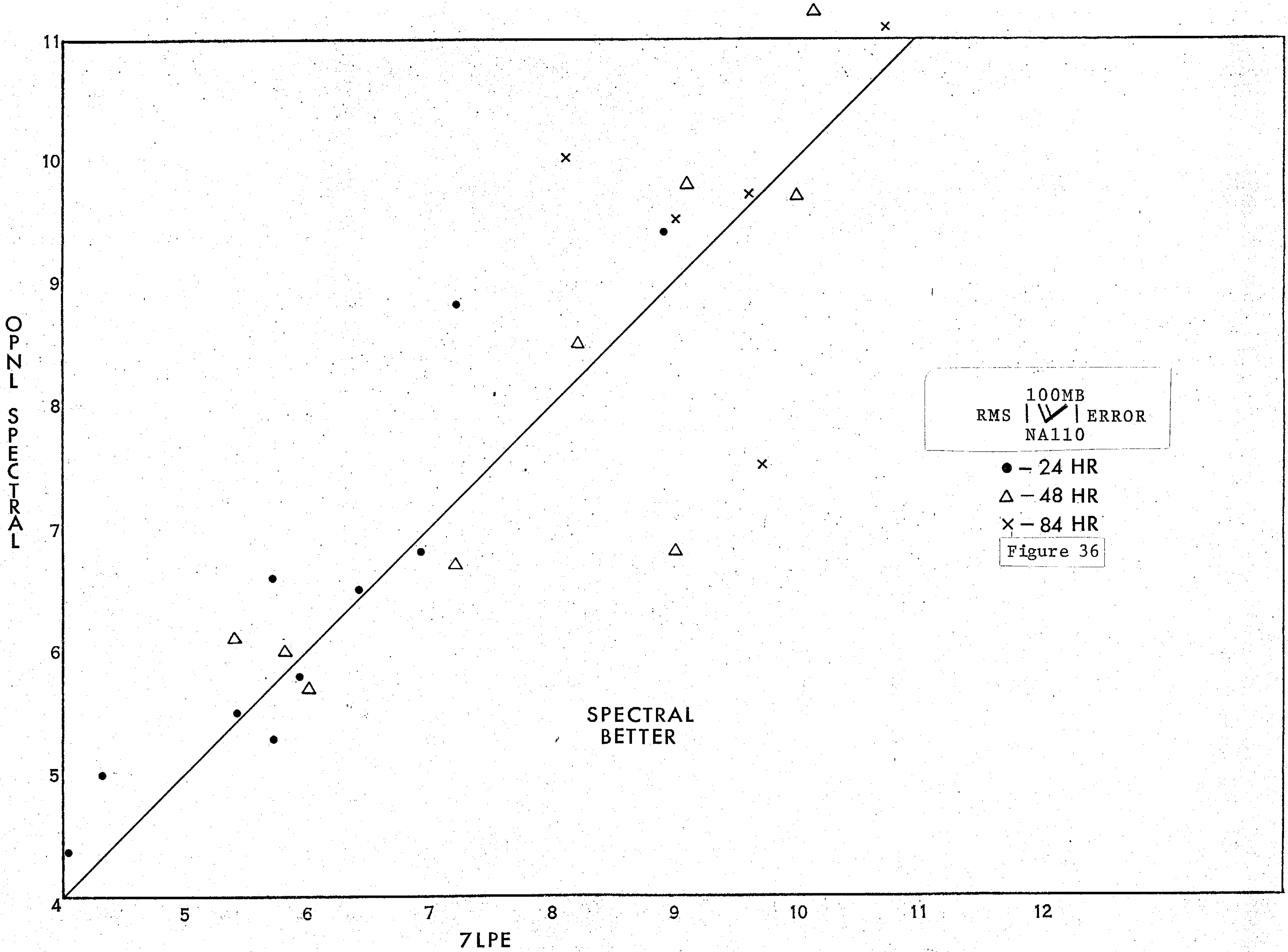
Figure 33



5.3
5.0
4.5
4.0
3.5
3.0
2.5
2.0
OPNL SPECTRAL
1.8
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
6.8
7LPE
SPECTRAL BETTER
100MB
RMS T ERROR
NA110
– 24 HR
– 48 HR
– 84 HR

Figure 34


100MB
MEAN SPEED ERROR
NA110
● — 24 HR
△ — 48 HR
× — 84 HR
Figure 35
SPECTRAL
BETTER
SPECTRAL
BETTER
OPNL SPECTRAL
7LPE
5
4
3
2
1
0
−1
−2
−4
−3
−2
−1
0
1
2
3
4
5



OPNL SPECTRAL
7LPE
4
5
6
7
8
9
10
11
12
SPECTRAL BETTER
100MB
RMS |V| ERROR
NA110
● – 24 HR
△ – 48 HR
× – 84 HR

Figure 36

SPECTRAL / 7LPE TEST : 12-48HRS (10 CASES), 60-84HRS (5 CASES)

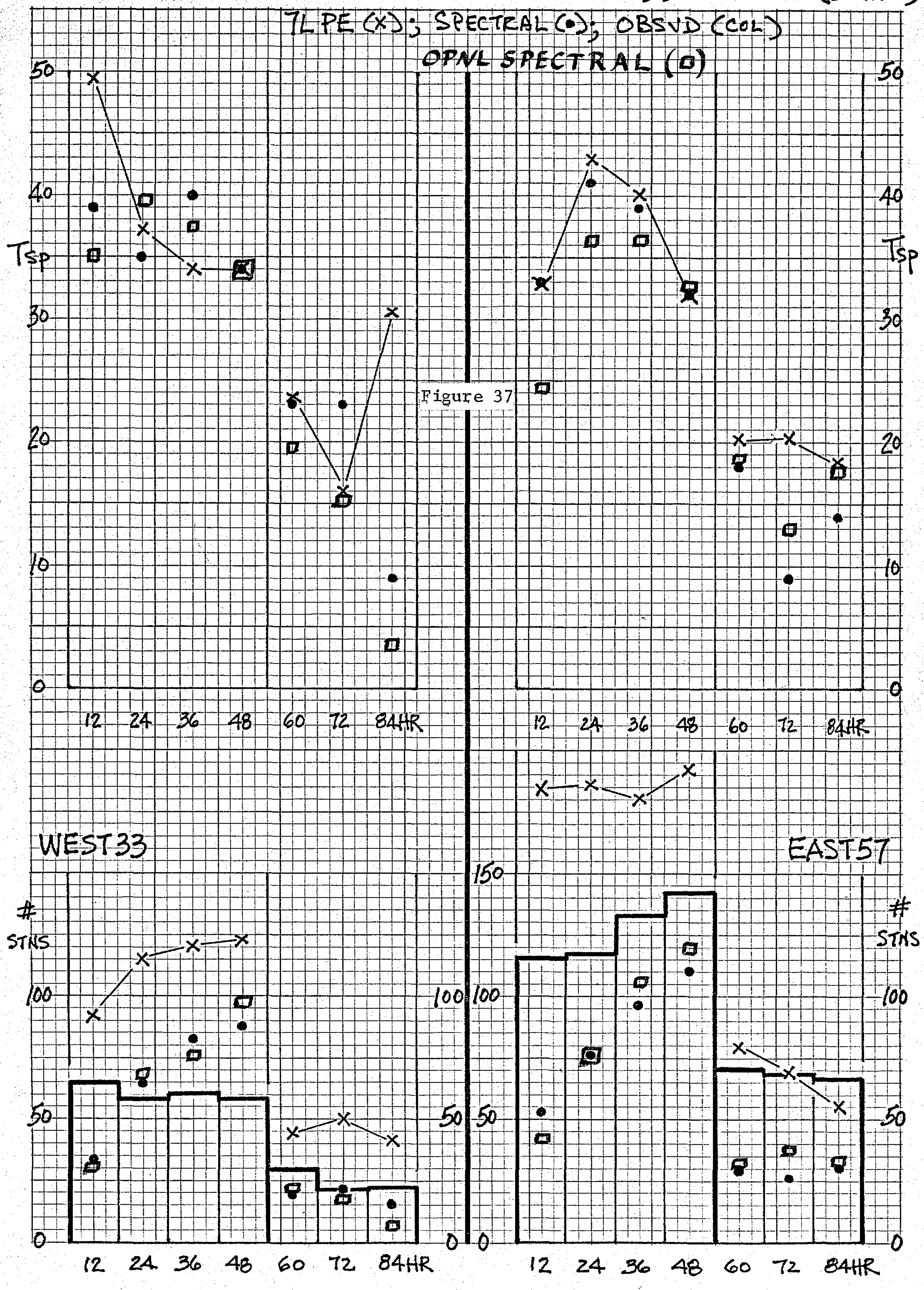


Figure 37

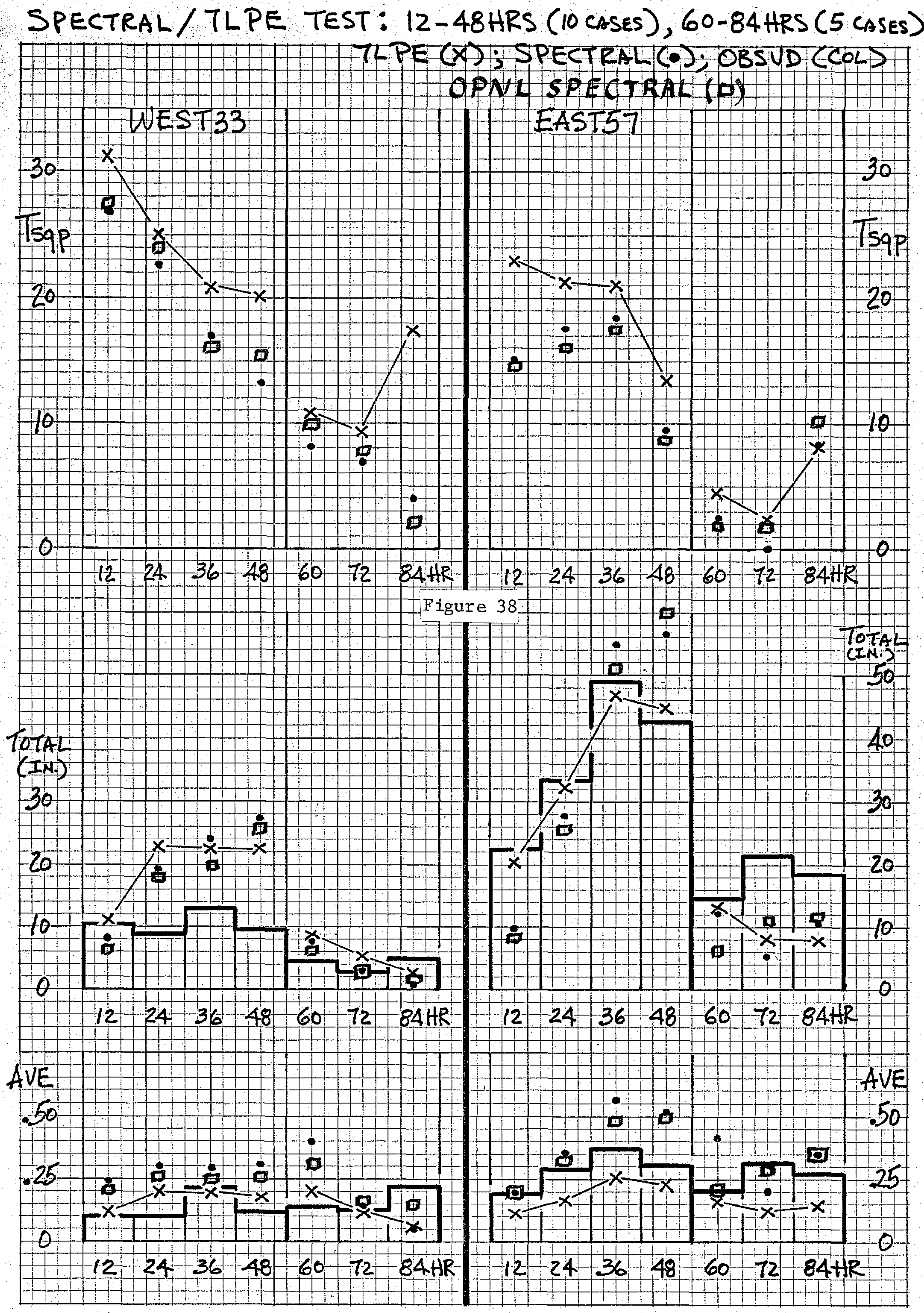

SPECTRAL / 7LPE TEST: 12-48HRS (10 CASES), 60-84HRS (5 CASES)
7LPE (X); SPECTRAL (•); OBSVD (COL)
OPNL SPECTRAL (▫)
WEST33
EAST57
Tsqp
30
20
10
0
12 24 36 48 60 72 84HR
TOTAL (IN.)
50
40
30
20
10
0
AVE
.50
.25
0

Figure 38

APPENDIX I

APPENDIX II

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Meteorological Center

APPENDIX II

May 1, 1980

TO : Spectral/7L PE Jury Member ________

FROM: OA/W324 - John D. Stackpole John D. Stackpole

SUBJ: Re-Evaluation of Spectral Model

As you know there were a few problems associated with the G3012 (Global 30 Mode 12 Layer) spectral model in the initial tests. The model has been slightly reconfigured (layers re-arranged, smoothing reduced, humidity initial conditions better specified) and we are in the process of re-running the test cases.

Here is the complete set of maps, with the new spectral (Labeled "SM G30/H24") forecast set on top, for case

______ Z 197 .

Note these are the originals (not xerox copies). PLEASE take care not to deface or lose any.

I am not asking for a complete re-assessment of the forecast quality, just your answer to this multiple choice question:

The New Spectral is:
- [ ] Better and has overcome my previous objections
- [ ] Little different
- [ ] Worse than before

(If you can't remember what your "previous objections" were let me know and I can dig out your earlier evaluation questionnaire). Comments Welcome ...

Please return this map bundle and this memo to Stackpole ASAP - we have to pass it on to the other jurors.

SUBJECTIVE EVALUATION FORM
7LPE vs. SPECTRAL

Instructions: Indicate by letter PE: 7LPE in the boxes
S: Spectral

which model forecast you think is better in your field of specialization only. Indicate a tie by a T and if one forecast is exceptionally better, add a "+" to the letter. i.e. "S" means Spectral better; "PE+" means 7LPE much better (there is no "slightly better" option).

Initial time/date of Forecast ____________________

Evaluated by: ____

Specialization Area

I. Maine, Alaska & West
(Harry Brown)

	24 hr.	48 hr.
SLP & Thickness Western Atlantic ---------------		
East Pacific & Alaska ----------		
500 ht & Vort. Western Atlantic ---------------		
East Pacific & Alaska ----------		

II. Aviation Forecasting
(Roy McCarter)

	24 hr.
250 mb Jet & Isotachs Atlantic -----------------------	
N. American --------------------	
Pacific ------------------------	
250 mb hts. & isotherms Atlantic ---------------------	
N. American --------------------	
Pacific -----------------------	

	24 hr.
Tropopause & Vert. Shear -------	
100 mb. hts. -------------------	
100 mb. isotachs ---------------	
100 mb. isotherms ---------------	
General Aviation needs ----------	

III. Quantitative Precipitation forecasting (Dave Olson)

	24 hr.	48 hr.	84 hr.
Rain/No Rain Coverage Eastern U. S. -------------------			
Western U. S. -------------------			
Quantitative Amounts Eastern U. S. -------------------			
Western U. S. -------------------			
Relative Humidity Patterns Eastern U. S. -------------------			
Westerns U. S. -----------------			
Utility of Mass & Wind Forecasts to QPF Forecasts U. S. -------------------------			

UNITED STATES DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
Rockville, Maryland 20852

November 28, 1979

APPENDIX I

TO : 7LPE vs. Spectral Jury Member ________

FROM: OA/W342 - John D. Stackpole JDStackpole

SUBJ: Evaluation of Relative Merits of 7LPE & Spectral Model in Your Field of Specialization

Here are the various maps you said earlier that you would need to form a subjective judgment on the relative merits of the two contending models.

You have here the forecasts and verifications for:

Case # ____________ Initial time ______ Z ______________

You also have here a fill the blanks questionnaire (with instructions) to allow you to express your opinions in a manner that can be tabulated with ease. The form has fill-in positions for all the specialists - please use only the portion that is yours.

When you have done your thing, please return the entire package to Stackpole in SEB; I shall pass them all on to Saylor who is doing a generalist evaluation of the cases.

Don't forget, you will be receiving 10 cases in all, probably at a rate of one or two per week for the next couple of months (as fast as the machine can crank them out). Please do your evaluations promptly so as to avoid backlogs - laggards will be hectored.

Attachments (maps, forms, etc.)

References

Brown, H., 1971: Teweles-Wobus S1 Score. NMC Technical Attachment #71-2, 1 June 1971, NMC, NWS, NOAA.

Campana, K. A., 1978: Real Data Experiments with a 4th Order Version of the Operational 7-Layer Model. NMC Office Note #188, NMC, NWS, NOAA.

Cooley, D. S. 1979 (a): Modifications of the Seven Layer LFM Model. Technical Procedures Bulletin #265, 17 April 1979, MSD, NWS, NOAA.

Cooley, D. S. 1979 (b): Assorted Changes in the 7L PE Model; Physics and Numerics Technical Procedures Bulletin #276, 16 May 1979, MSD, NWS, NOAA.

Cressman, G. P., 1980: A Cautionary Note on the Use of N. M. C. Vertical Velocity Fields. NMC Office Note #213, NMC, NWS, NOAA.

Sela, J., 1980: Spectral Modeling at NMC, MS submitted to Monthly Weather Review.

Shuman, F. G., J. B. Hovermale, 1968: An Operational Six-Layer Primitive Equation Forecast Model. J. Appl. Meteor., 7, pp 525-547.

Stackpole, J. D. et al, 1978: How to Pick a New Forecast Model, Preprint Volume; Conference on Weather Prediction and Analysis and Aviation Meteorology, AMS Boston, MA., October 1978.

Stackpole, J. D., 1978: The National Meteorological Center's Seven Layer Model on a Northern Hemisphere 190.5 km Grid. NMC Office Note #177, NMC, NWS, NOAA.

Stackpole, J. D., 1978: The NMC 9 Layer Global Primitive Equation Model on a Latitude-Longitude Grid. NMC Office Note #178, NMC, NWS, NOAA.